



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 190359**

**TO: Celine Qian**  
**Location: rem/2A64/2C70**  
**Art Unit: 1636**  
**Friday, June 09, 2006**  
**Case Serial Number: 10/099663**

**From: Kristine Hensle**  
**Location: Biotech-Chem Library**  
**REM-1B69**  
**Phone: (571) 272-4161**

**Kristine. Hensle@uspto.gov**

### **Search Notes**

Examiner Qian,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Kristine Hensle  
Librarian (ASRC Aerospace)  
STIC Biotech/Chem Library  
(571)272-4161

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: May 28, 2006, 05:53:57 ; Search time 13877.8 Seconds  
(without alignments)  
10971.408 Million cell updates/sec

Title: US-10-099-663-1

Perfect score: 2381  
Sequence: 1 agcttcctgcgcagaaag.....atctgaagctccattcag 2381

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_env:\*  
2: gb\_pat:\*  
3: gb\_ph:\*  
4: gb\_dl:\*  
5: gb\_pr:\*  
6: gb\_ro:\*  
7: gb\_rts:\*  
8: gb\_sy:\*  
9: gb\_un:\*  
10: gb\_vl:\*  
11: gb\_ov:\*  
12: gb\_htg:\*  
13: gb\_in:\*  
14: gb\_om:\*  
15: gb\_ba:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	736.8	30.9	2847	11 AY254202	AY254202 Gallus ga
2	85	3.6	923	11 BX932340	BX932340 Gallus ga
3	81	3.4	899	11 BX930518	BX930518 Gallus ga
4	78.8	3.3	896	11 BX933597	BX933597 Gallus ga
5	64.2	2.7	1141	2 AR579680	AR579680 Sequence
6	64.2	2.7	1141	2 AX083744	AX083744 Sequence
7	62	2.6	3013	6 RATFABP	M18080 Rat Intesti
8	62	2.6	230897	12 AC130496	AC130496 Rattus no
9	60.2	2.5	320940	12 AC152772	AC152772 Bos tauru
10	57.8	2.4	1141	2 AR579680	AR579680 Sequence
11	57.8	2.4	1141	2 AX083744	AX083744 Sequence
12	57.6	2.4	320940	12 AC152772	AC152772 Bos tauru
13	56.6	2.4	5039	6 MUSFABP1	M65033 Mouse Fabd1
14	56.6	2.4	188707	12 AC158770	AC158770 Mus muscu
15	56.6	2.4	240767	6 AC155158	AC155158 Mus muscu
16	56.2	2.4	3509	14 DQ126268	DQ126268 Sus scrofa
17	55	2.3	713	14 AY911349	AY911349 Bos tauru
18	51.8	2.2	234081	13 PFMAL4P2	AL035475 Plasmodu

19	50.2	2.1	614	14 AY60624	AY60624 Sus scrofa
20	50.2	2.1	1522	13 AY701231	AY701231 Orconecte
21	50	2.1	564	6 RATFABPX	M35992 Rat Intesti
22	50	2.1	187264	12 CR847945	CR847945 Danio rer
23	49.8	2.1	619	11 XELIFABP	L19946 Xenopus lae
24	49.8	2.1	5204	2 AY771590	AY771590 Sequence
25	49.8	2.1	5204	5 HOMFABP	M18079 Human, inte
26	49.8	2.1	200000	12 AC008077	AC008077 Homo sapi
27	49.2	2.1	426	2 CS177440	CS177440 Sequence
28	49.2	2.1	2000	2 AX655393	AX655393 Sequence
29	49	2.1	134124	6 AC154837	AC154837 Mus muscu
30	49	2.1	202804	6 AC021630	AC021630 Mus muscu
31	49	2.1	249943	13 AB014823	AB014823 Plasmodu
32	48.6	2.0	176386	12 CR855271	CR855271 Danio rer
33	48.4	2.0	564	6 RATFABP1	K01180 Rat Intesti
34	48.4	2.0	832	2 AX972010	AX972010 Sequence
35	48.4	2.0	832	2 BD110729	BD110729 EST and e
36	48.4	2.0	832	2 AR415176	AR415176 Sequence
37	48.4	2.0	18876	13 CER3208	Z72509 Caenorhabdi
38	48.4	2.0	151090	11 AL929044	AL929044 Zebrafish
39	48.4	2.0	176472	11 CR762472	CR762472 Zebrafish
40	48.4	2.0	197110	5 AC104306	AC104306 Homo sapi
41	48.2	2.0	132254	13 AC116330	AC116330 Dictyoste
42	48.2	2.0	132254	13 AC116330	AC116330 Dictyoste
43	48	2.0	191840	5 AC092656	AC092656 Homo sapi
44	47.8	2.0	3853	11 AY266452	AY266452 Danio rer
45	47.8	2.0	110000	13 AC116984_1	Continuation (2 of

## ALIGNMENTS

RESULT 1	AY254202	2847 bp	DNA	linear	VRT 22-APR-2003
LOCUS	AY254202				
DEFINITION	Gallus gallus intestinal fatty acid-binding protein gene, complete cds.				
ACCESSION	AY254202.1	GI:30060211			
VERSION	AY254202.1	GI:30060211			
KEYWORDS					
SOURCE	Gallus gallus (chicken)				
ORGANISM	Gallus gallus				
REFERENCE	1 (bases 1 to 2847) Wang, Q., Li, H., Wang, Y. and Zhao, J. Cloning and characterization of chicken I-FABP gene Unpublished				
AUTHORS	2 (bases 1 to 2847) Wang, Q., Li, H., Wang, Y. and Zhao, J. Direct Submission Submitted (11-MAR-2003) Animal Science & Technology, Northeast Agricultural University, 59 Mucal Street, Harbin, Heilongjiang 150030, China				
TITLE	Location/Qualifiers				
JOURNAL	1. .2847 /organism="Gallus gallus" /mol_type="genomic DNA" /db_xref="taxon:9031" /rname="blood" /note="Breed: Arber Acres broiler" join(21. .87,720. .892,1465. .1572,2416. .2466) product="intestinal fatty acid-binding protein" join(21. .87,720. .892,1465. .1572,2416. .2466) /note="FABP" /codon_start=1 /product="intestinal fatty acid-binding protein" /protein_id="AA013101.1" /db_xref="GI:30060212" /translation="MANGTWKIEKKNENYKPEAMGVNVMKRLGADNDUKLTIOOD GNKFLVSSNFRITDIEFLGVSPFSLADGTELSGWNLEGNKVLGTETRDNGRV LTAYREIVGSELIQTYVEGVENKRIFFKE"				
FEATURES	source				
CDS					
mRNA					

## ORIGIN

Query Match 30.9%; Score 736.8; DB 11; Length 2847;  
 Best Local Similarity 98.1%; Pred. No. 8.9e-221;  
 Matches 757; Conservative 0; Mismatches 12; Indels 3; Gaps 1;

QY 1613 CTACAGACAGAAAGATGGCGATTTAAGGATCTTGGAAAAATAGAAAAATGAACTATG 1672  
 DB 7 CGAGAGACAGAAAGATGGCGATTTAAGGATCTTGGAAAAATGAACTATG 66  
 QY 1673 AAAAATTCATGAGCAATGGGTAAGCCTTCTTTTGAATGCCCTTCTTAAAGAGAGAT 1732  
 DB 67 AAAAATTCATGAGCAATGGGTAAGCCTTCTTTTGAATGCCCTTCTTAAAGAGAGAT 126  
 QY 1733 ACCACATGCGCGGAAATACAACTTAGCGTTCAAGAACTACCATCTGCTAAGCTGTC 1792  
 DB 127 ACCACATGCGCGGAAATACAACTTAGCGTTCAAGAACTACCATCTGCTAAGCTGTC 186  
 QY 1793 TTTGTTGCTGCTAATTTTGCCTTGACATTCGCCCTGCACTTATTTTGAAGACCTAT 1852  
 DB 187 TTTGTTGCTGCTAATTTTGCCTTGACATTCGCCCTGCACTTATTTTGAAGACCTAT 246  
 QY 1853 AAGAGGGAATACAGAGAAAAACATTCGATTTTATTTGATTCGATTAATCTTATGC 1912  
 DB 247 AAGAGGGAATACAGAGAAAAACATTCGATTTTATTTGATTCGATTAATCTTATGC 306  
 QY 1913 ATTAGCTAATTCAGTAGAGGATTCAGAGAAATTTAATAGATTAATATATGATGGA 1972  
 DB 307 ATTAGCTAATTCAGTAGAGGATTCAGAGAAATTTAATAGATTAATATATGATGGA 366  
 QY 1973 ATATTATTTTGAATGACTGTTGAAAAATTAACAGAGGGAATTCCTGCTCCAGT 2032  
 DB 367 ATATTATTTTGAATGACTGTTGAAAAATTAACAGAGGGAATTCCTGCTCCAGT 426  
 QY 2033 TTTGAGAGACACACATGATTTT---GAGTCATTTTAAAGTCTAGGCTTACTTTAACTT 2089  
 DB 427 TTTGAGAGACACACATGATTTTGAATCAATTTTAAAGTCTAGGCTTACTTTAACTT 486  
 QY 2089 GTACAAATCTGCTGTAATATGATGATTAACATTAATCTTATGTTGATGATGTTGTA 2149  
 DB 487 GTACAAATCTGCTGTAATATGATGATTAACATTAATCTTATGTTGATGATGTTGTA 546  
 QY 2150 TTACAGGCTGAACCTGCTCAGTGAAGGATGGAAGAGATTAAGCTCTGAGTCAGAAT 2209  
 DB 547 TTACAGGCTGAACCTGCTCAGTGAAGGATGGAAGAGATTAAGCTCTGAGTCAGAAT 606  
 QY 2210 CTGGGCTAAGCTCCCTCAACTACAGAAAAAGTCAATTAATAAATGCAACATGATGTC 2269  
 DB 607 CTGGGCTAAGCTCCCTCAAGTACAGAAAAAGTCAATTAATAAATGCAACATGATGTC 666  
 QY 2270 ATTTGTTTTTCTCTGCTGATGTTAATGATTAATTAATTTTATTTTATAGCGGTGA 2329  
 DB 667 ATTTGTTTTTCTCTGCTGATGTTAATGATTAATTAATTTTATTTTATAGCGGTGA 726  
 QY 2330 TGTGATGAAAAAGAAAGTTAGAGCCGACGATTAATCTGAAGCTCACTATTCAG 2381  
 DB 727 TGTGATGAAAAAGAAAGTTAGAGCCGACGATTAATCTGAAGCTCACTATTCAG 778

RESULT 2  
 LOCUS BX932340 923 bp mRNA linear VRT 30-MAR-2004  
 DEFINITION Gallus gallus finished cDNA, clone CHEST35022.  
 ACCESSION BX932340  
 VERSION BX932340.2 GI:46017395  
 KEYWORDS  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 REFERENCE 1 (bases 1 to 923)  
 AUTHORS Boardman, P.E., Bonfield, J.K., Brown, W.R.A., Carder, C., Chalk, S.E.,

TITLE  
 JOURNAL  
 COMMENT  
 Submitted (29-MAR-2004) Sanger Institute, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: chickens@ems.umbc.ac.uk  
 On Apr 1, 2004 this sequence version replaced gi:41632868.  
 BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cDNA  
 sequencing project.  
 This sequence is from the  
 BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,  
 from a library constructed by Elizabeth Bosch. cDNA was prepared  
 from RNA extracted from heart, normalised, and poly A-tailed.  
 EcoRI-NcoI cut cDNA was then ligated into the vector. Vector:  
 pBluescript II KS(+); Site\_1: EcoRI; Site\_2: NcoI Host: Escherichia  
 coli DH10B.  
 FEATURES  
 source location/Qualifiers  
 1..923  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /strain="White Leghorn, Hisex"  
 /db\_xref="taxon:9031"  
 /clone="CHEST35022"  
 /clone\_11b="CSECHN70"  
 /dev\_stage="stage 36"

## ORIGIN

Query Match 3.6%; Score 85; DB 11; Length 923;  
 Best Local Similarity 100.0%; Pred. No. 1e-14;  
 Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1610 TGGCTTACAGACAGAAAGATGGCGATTTAAGGATCTTGGAAAAATAGAAAAATGAACT 1669  
 DB 1 TGGCTTACAGACAGAAAGATGGCGATTTAAGGATCTTGGAAAAATAGAAAAATGAACT 60  
 QY 1670 ATGAAAATTCATGAGCAATGGG 1694  
 DB 61 ATGAAAATTCATGAGCAATGGG 85

RESULT 3  
 LOCUS BX930518 899 bp mRNA linear VRT 02-FEB-2004  
 DEFINITION Gallus gallus finished cDNA, clone CHEST640b17.  
 ACCESSION BX930518  
 VERSION BX930518.1 GI:41631046  
 KEYWORDS  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 REFERENCE 1 (bases 1 to 899)  
 AUTHORS Boardman, P.E., Bonfield, J.K., Brown, W.R.A., Carder, C., Chalk, S.E.,  
 Croxson, M.D.R., Davies, R.M., Francis, M.D., Grafham, D.V.,  
 Hubbard, S.J., Humphray, S.J., Hunt, P.J., Maddison, M., McLaren, S.R.,  
 Niblett, D., Overton, I.M., Rogers, J., Scott, C.E., Taylor, R.G.,  
 Tickle, C. and Wilson, S.A.  
 Direct Submission  
 Submitted (02-FEB-2004) Sanger Institute, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: chickens@ems.umbc.ac.uk  
 BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cDNA  
 sequencing project.  
 This sequence is from the  
 BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,  
 from a library constructed by Elizabeth Bosch. cDNA was prepared  
 from RNA extracted from small intestine, normalised, and poly  
 A-tailed. EcoRI-NcoI cut cDNA was then ligated into the vector.  
 Vector: pBluescript II KS(+); Site\_1: EcoRI; Site\_2: NcoI Host:  
 Escherichia coli DH10B.  
 FEATURES  
 source location/Qualifiers  
 1..899



Dy 148I GTGACCCCCCTCCTTGACAAGATCAATGCCAAGTTGAAGCTTAAAGCCAGGCATCATCA 1540A  
Dd :  
Db NNDAGAMDWITYMGNGNTMMERRAMNNMMAICBPAYCCNNNNNPAVCVHKHKKYMRITWK 410  
Dy 154I TGTAAATTGCTTCCTCGATAGCCGGTTCATAATTCCTTTGCCAAAGCTGCTCATTA 1600A  
Dd :  
Db YMKKACANNNNBKAKMYRMVAMMYISBDITINDIMMMWTSDMBWHMTYVTYTVMRANNNNN 350  
Dy 160I CCAGAAGTCTGCTTACAGACAGAAAAGATGCGATTTCAGGTACTTGGAAAAATAGAGAAA 1660A  
Dd :  
Dy 349 NNWRCKTTSWMNMMDHMNHICTYGNNTWSAYSBAAMSMAAGASNBVTYNMCRRITYM 290  
Dd :  
Dy 166I ATGAAACTATGAAAAAT-TCATGGAACCAATGGGTAAAGCTTACTTTTTGAATGCTT 1719A  
Dd :  
Dy 289 GKTIHTNNNNNNKAMYRTKTVAJCNRRYYUPTAVTBGRNKYCYPABWYBYBMVGZHHV 230  
Dd :  
Dy 172O CTAAGAAGAGATACCACCTAGCGGSGAATACAACTTAGCTGTTCATGAACTACATCT 1779A  
Dd :  
Dy 229 BWMRBABRBSWNMMVVVKCRNKYUVSMYHAMRYBRKABAYOGCNNMMXDRRAHHHCATNN 170  
Dd :  
Dy 178O GGCCTAACCTGCTTTGTGTGTCTCTATTTTGGCCCTTGCAATTCGCCCTGCATTTATTT 1839A  
Dd :  
Dy 169 NMNMNVAVYMHMKKGDAWTNNKTJBDRDBHBHVTTYWVRDYWCAMCMNAKAKVR 110  
Dd :  
Dy 184O GAAGAAGCTCATAGAGGGGATAACAAGAAAGAAAACATTCGATTTATTTTGC 1894A  
Dd :  
Dy 109 TAMGHMYTYDRYVSANTGVRRMMHMCSCMYSTNNNRMTYRNKRXTVAMAMTM 55

RESULT 6			
AX083744/C			
LOCUS	AX083744	1141 bp	DNA
DEFINITION	Sequence 22 from Patent WO0111061.		linear
ACCESSION	AX083744		
VERSION	AX083744.1	GI:13185472	
KEYWORDS			
SOURCE			
ORGANISM	synthetic construct		
	synthetic construct		
REFERENCE	1	other sequences; artificial sequences.	
AUTHORS	Kunst, L. and Clemens, S.		
TITLE	Regulation of embryonic transcription in plants		
JOURNAL	Patent: WO 0111061-A 22.15.FEB-2001;		
	UNIVERSITY OF BRITISH COLUMBIA (CA)		
FEATURES	Location/Qualifiers		
Source	1..1141		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
	1..1141		
	/note="Consensus sequence of A.t., L.a., and B.n. FABI		
	promoters"		
PROMOTER			
promoter			
ORIGIN			

Query Match	2.7%	Score 64.2	DB 2	Length 1141
Best Local Similarity	12.7%	Pred. No. 3.76e-08		
Matches	136	Conservative 379	Mismatches 550	Indels 10
			Gaps	3
QY	830	TTTTCTGCATTATTCCTTCAACATTAAACCTGGAGCTATGAGAAATCAACACGTTGGG	889	
DB	1129	TTTCTTATTKYKANNNNNNNNNGMKGNRRDATTMSATGTAMTNNNAKRGATCMCTWYTTGTGN	1070	
QY	890	TAAATTCGACTAGACAGACATCAACTACGTGTAGAAATGACAGAAACAGACATTCCT	949	
DB	1069	RMCMCHRTYAMRTWYTRSNAAWSGCTKEMWMTMKRYATKRTATYAAWCAAMRRNNWCACTN	1010	
QY	950	GAATGGGCTATTAATTTAGAGAAATTCGTAAGGTGTCTGAAATTTAGTACTCTATTAA	1009	
DB	1009	GAASCATNNMMWYATTMAAYAAAKAPAGAGNNMEMYGAAGNKKMGAAANATGEBWMA	950	
QY	1010	GAGTGAGACACGAGATGAGAAATATATCGCAATTTCTTACTCTAGACACTAGACTCGAA	1069	
DB	949	DTAGMCCNNNNNNNTTVRRMAKAKAKNNNNNNATTAACNTRATNNKAAATTHMKRTTGDAH	890	

[illegible]

RESULT 7	REFABP	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
	RATFPAB		3013 bp	DNA	linear	ROD 27-APR-199				
						Rat intestinal fatty acid binding protein gene, exons 1 and 2, and a B2 repetitive element.				
						M18080 J03465				
						M18080.1 GI:204071				
						B2 repetitive sequence; fatty acid binding protein.				
						Rattus norvegicus (Norway rat)				
						Rattus norvegicus				
						Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Rattus.				
						1 (bases 1 to 3013)				
						Sweetser,D.				

JOURNAL Unpublished (1987)  
 REFERENCE 2 (bases 917 to 1217)  
 AUTHORS Sweetser,D.A., Birkenmeier,E.H., Klieak,I.J., Zollman,S.,  
 Sparke,R.S., Mohandas,T., Lueis,A.J. and Gordon,J.I.  
 TITLE The human and rodent intestinal fatty acid binding protein genes. A  
 comparative analysis of their structure, expression, and linkage  
 relationships  
 JOURNAL J. Biol. Chem. 262 (33), 16060-16071 (1987)  
 PUBMED 2824476  
 COMMENT Original source text: Rat (Sprague Dawley) intestinal epithelial  
 lining DNA, clone lambda-RFAPB.  
 Draft entry and computer-readable sequence for [2] kindly provided  
 by D.Sweetser, 19-JAN-1988.  
 FEATURES  
 source location/Qualifiers  
 1..3013  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 repeat\_region 195..406  
 /note="B2 repeat"  
 prim\_transcript 1155..>3013  
 /note="FABP mRNA (alt.) and introns"  
 prim\_transcript 1179..>3013  
 /note="FABP mRNA (alt.) and introns"  
 CDS join(1215..1281,2541..2713)  
 /note="fatty acid binding protein"  
 /codon\_start=1  
 /protein\_id="AA041133.1"  
 /db\_xref="GI:204072"  
 /translation="MAFDGTMKVDNENYKEMRMGINVVRKLGANDLKLTTTQ  
 GNMPTVKESSNFRNDVVEELGVDPAYSLADGTBLT"  
 <1215..1281  
 /note="fatty acid binding protein"  
 /number=1  
 intron 1282..2540  
 /note="FABP intron A"  
 2541..>2713  
 /note="fatty acid binding protein"  
 /number=2  
 intron 2714..>3013  
 /note="FABP intron B"  
 ORIGIN 1 bp upstream of EcoRI site.  
 Query Match 2.6%; Score 62; DB 6; Length 3013;  
 Best Local Similarity 59.3%; Pred. No. 1.9e-07;  
 Matches 127; Conservative 0; Mismatches 80; Indels 7; Gaps 1;  
 QY 1515 TTGAGCTTTCACGACCATCATCATTAATGCTTCTGATTAAGCTTTGATTA  
 1574  
 Db 1096 TTGAACCTTGAACCTTCACATCATCATGATGTTGCGAAGATTAAGAAATTA  
 1155  
 QY 1575 TTCTCTTTCGAAAGCTCTGCTACTTACGAAAGTC-----TGCTTCACAGAGA  
 1627  
 Db 1156 TTCTCTTTCGAGCAGAGCAGAACTCTGCTTTCCTAGAGGCACACAGCTGACATCA  
 1215  
 QY 1628 TGGCATTTAAGCTTTCGAAATAATGAGAAATGAGAAATTCATGAGAG 1687  
 Db 1216 TGGCATTTGATGACCTTGAAGAGTACCGGAATGAGAAATGATTAAGTTCAATGAGA  
 1275  
 QY 1688 CAATGGGTAGCCTTACTTTTTCGAAATGCTTCT 1721  
 Db 1276 AATGGGTAGGCGCTCTCTGCTATTTGCT 1309

ORGANISM  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Muridea; Muridae; Murinae; Rattus.  
 REFERENCE 1 (bases 1 to 230897)  
 AUTHORS Muzny,D.,Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,  
 Allen,C., Allen,H., Alsbrooks,S., Amn,A., Angiulo,D.,  
 Anyalebech,V., Ayogei,A., Ayodeji,M., Baca,E., Baden,H.,  
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,  
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,  
 Bryant,N., Buhay,C., Burch,P., Butrell,K., Calderon,E.,  
 Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,  
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,  
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,  
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,  
 Delgado,O., Denson,S., Dermo,C., Ding,Y., Dinh,H., Divya,K.,  
 Drapep,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,  
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,  
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,  
 Frazer,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,  
 Gebregeorgis,B., Geer,K., Gill,R., Grady,M., Guerra,W., Guvera,W.,  
 Gunaratne,P., Healand,W., Hamil,C., Hamilton,C., Hamilton,K.,  
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,  
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,W.,  
 Hollins,B., Howells,S., Huylk,S., Hume,J., Idlebird,D., Jackson,A.,  
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolyet,A.,  
 Karpathy,S., Kelly,S., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,  
 Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,  
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,  
 Lotenshewa,L., Louised,H., Lozada,M., Malmoud,M., Malloy,K., Mangum,A.,  
 Maneshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,  
 Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,  
 Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,  
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,  
 Morgan,M., Morris,K., Morris,S., Mundaya,M., Murphy,M., Natl,L.,  
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,  
 Nwokoelamoh,O., Okunou,G., Olariunsaogun,A., Pal,S., Parks,K.,  
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,  
 Plopper,F., Poindecker,A., Popovic,D., Prims,E., Pu L.-L.,  
 Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,  
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Rhodes,S., Riggs,F.,  
 Rivers,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,  
 Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,  
 Shetty,J., Shvartsbeyn,A., Slason,I., Sitter,C.D., Smajz,D.,  
 Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,  
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,  
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Tjoej,Z., Uman,K.,  
 Valas,R., Vera,V., Villalana,D., Waldron,L., Walker,B., Wang,J.,  
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,  
 Williams,G., Willson,R., Wleczky,R., Wooden,H., Morley,K.,  
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,  
 Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von  
 Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,  
 Weinstein,G. and Gibbs,R.A.  
 TITLE Direct Submission  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 230897)  
 AUTHORS Morley,K.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (09-NOV-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 QY 3 (bases 1 to 230897)  
 Db Rat Genome Sequencing Consortium.  
 COMMENT Direct Submission  
 Submitted (09-NOV-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Nov 9, 2002 this sequence version replaced gi:23101530.  
 The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
 in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

## ----- Genome Center

Center: Baylor College of Medicine  
Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu

## ----- Project Information

Center project name: GBMN  
Center clone name: CH230-4G2

## ----- Summary Statistics

Assembly program: Phrap; version 0.990329  
Consensus quality: 210222 bases at least Q40  
Consensus quality: 213326 bases at least Q30  
Consensus quality: 215519 bases at least Q20  
Estimated insert size: 218065; sum-of-contigs estimation  
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

## ----- NOTE: Estimated insert size may differ from sequence length

(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).

NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 157719: contig of 157719 bp in length  
\* 157720 157819: gap of unknown length  
\* 157820 230897: contig of 73078 bp in length.

## Location/Qualifiers

1. 230897  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-4G2"

## misc\_feature

1. 2206

/note="wgs\_end\_extension  
clone\_end:Sp6  
complement(72249..73061)  
/note="clone\_boundary  
clone\_end:Sp6  
site:EcORI

## misc\_feature

end\_sequence.BH306439"  
157720..157819  
/estimated\_length=unknown  
157820..159219  
/note="wgs\_contig"  
182442..183960  
/note="wgs\_contig"

## misc\_feature

157720..157819  
/estimated\_length=unknown  
157820..159219  
/note="wgs\_contig"  
182442..183960  
/note="wgs\_contig"

## ORIGIN

## Query Match

Best Local Similarity 59.3%; Pred. No. 1.9e-07;  
Matches 127; Conservative 0; Mismatches 80; Indels 7; Gaps 1;

## ORIGIN

1515 TTGAGCTTTAGCCACGACATCATGTAATGCTTTCTGATTAAGCTGTTCAATAA 1574  
|||||  
156173 TTGAACTTTGAACCTTCACATCATGTAATGATTTGCGAAGATTAAGAAATTAATAA 156232  
|||||

## ORIGIN

1575 TTCTCTTGCAGAGCTCTGCTACTTACCAAGAGC-----TGCCTACAGACAGGAAGA 1627  
|||||  
156233 TTCTCTTGTAGTGACAGGACCGAAATCTGCTTTCTTAGAGGCAACACACAGCTGACATCA 156292  
|||||

## ORIGIN

1628 TGGCATTTAAAGGTAAGTAAGAAATAGAAAGAACTATGAATAATTCATGAGAG 1687  
|||||

Db 156293 TGGCATTTGATGACACTTGGAAAGTAGACCGAATGAACTATGAATAAGTTCAATGAGA 156352  
Oy 1688 CAATGGTAGACCTTACTTTTGAATGCCCTCT 1721  
|||  
Db 156353 AAATGGTAGAGGCTGGCTTCTGCTATTGCT 156386  
|||

## RESULT 9

## AC152772

## LOCUS

AC152772 320940 bp DNA linear HTG 23-JAN-2005  
Bos taurus clone CH240-10M22, WORKING DRAFT SEQUENCE, 81 unordered  
pieces.

## ACCESSION

AC152772

## VERSION

AC152772.2 GI:58037730

## KEYWORDS

HTG; HTGS PHASE1; HTGS\_DRAFT.

## SOURCE

Bos taurus (cattle)

## ORGANISM

## REFERENCE

## AUTHORS

Mizny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
Aisbrooks,S.L., Amaralunga,H.C., Ate,J.R., Ayala,M., Banke,T.,  
Barbata,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,  
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dahorne,S.R., David,R.,  
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
DeLaney,K.R., Delgado,O., Denu,A.L., Ding,Y., Dinb,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Earmhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falle,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Franz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Gara,N., Gill,R.,  
Gorell,J.H., Guenara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hogue,M., Hollway,C., Hollins,B.,  
Homesi,F., Howard,S., Huber,J., Hulik,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlsom,E., Kelly,S., Khan,U., King,L., Korrah,J., Kovar,C.,  
Kratovic,J., Kuirehl,A., Landry,N., Leal,B., Lewis,J.C., Lewis,L.,  
Li,J., Li,Z., Lichtharge,O., Lieu,C., Liu,J., Liu,W., Louleghed,H.,  
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Mosier,M., Neal,D., Newton,J., Newson,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okuwonu,G.,  
Oragunye,N., Oyiedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Prims,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,  
Schere,S., Scott,G., Shen,H., Shoshitari,N., Slason,I.,  
Sodergren,E., Sonake,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Umanai,K., Vasquez,L., Vera,V., Verdusco,D., Villalon,D.,  
Vinson,R., Waldron,L.M., Wang,Q., Wang,S., Ward-Moore,S.,  
Warren,R., Washington,C., Watlington,S., Williams,G.,  
Williamson,A., Wlezyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,  
Wu,Y.F., Zhou,D., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.

## TITLE

## JOURNAL

## AUTHORS

## JOURNAL

## REFERENCE

## AUTHORS

## JOURNAL

Submitted (23-JAN-2005) Human Genome Sequencing Center, Department



## COMMENT

of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Jan 23, 2005 this sequence version replaced gi:55819826.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

## ----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

## ----- Project Information

Center project name: PATP

Center clone name: CH240-10M22

## ----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 288280 bases at least Q40

Consensus quality: 294971 bases at least Q30

Consensus quality: 301063 bases at least Q20

Estimated insert size: 229238; sum-of-contigs estimation

Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

## -----

NOTE: Estimated insert size may differ from sequence length

(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html))

NOTE: This sequence may represent more than one clone.

NOTE: This is a 'working draft' sequence. It currently

consists of 81 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 5541: contig of 5541 bp in length  
5542 5591: gap of 50 bp  
5592 9604: contig of 4013 bp in length  
9605 9705: gap of unknown length  
9705 11503: contig of 1799 bp in length  
11504 11802: gap of 299 bp  
11803 20154: contig of 8352 bp in length  
20155 20254: gap of unknown length  
20255 24517: contig of 4263 bp in length  
24518 24567: gap of 50 bp  
24568 25759: contig of 1192 bp in length  
25760 26270: gap of 511 bp  
26271 29359: contig of 3089 bp in length  
29360 29459: gap of unknown length  
29460 32346: contig of 2886 bp in length  
32346 32395: gap of 50 bp  
32396 38577: contig of 6182 bp in length  
38578 38677: gap of unknown length  
38678 40965: contig of 2288 bp in length  
40966 41015: gap of 50 bp  
41016 47383: contig of 6368 bp in length  
47384 47484: gap of unknown length  
47485 48762: contig of 1278 bp in length  
48762 48843: gap of 81 bp  
48843 51170: contig of 2328 bp in length  
51171 51270: gap of unknown length  
51271 53092: contig of 1822 bp in length  
53093 53142: gap of 50 bp  
53143 55224: contig of 2082 bp in length  
55225 55313: gap of 89 bp  
55314 57275: contig of 1962 bp in length

57276 57375: gap of unknown length  
57376 63877: contig of 6502 bp in length  
63878 63927: gap of 50 bp  
63928 65212: contig of 1285 bp in length  
65213 65262: gap of 50 bp  
65263 66517: contig of 1555 bp in length  
66518 66617: gap of unknown length  
66618 70198: contig of 3581 bp in length  
70199 71339: gap of 1141 bp  
71340 75208: contig of 3669 bp in length  
75209 75308: gap of unknown length  
75309 97232: contig of 21924 bp in length  
97233 97282: gap of 50 bp  
97283 100331: contig of 3949 bp in length  
100332 101076: gap of 745 bp  
101077 102495: contig of 1419 bp in length  
102496 102595: gap of unknown length  
102596 104484: contig of 1689 bp in length  
104485 104708: gap of 224 bp  
104709 109824: contig of 5116 bp in length  
109825 115197: contig of 5273 bp in length  
115198 115247: gap of 50 bp  
115248 123071: contig of 7824 bp in length  
123072 123171: gap of unknown length  
123172 131072: contig of 7901 bp in length  
131073 131122: gap of 50 bp  
131123 135372: contig of 4250 bp in length  
135373 135472: gap of unknown length  
135473 140453: contig of 4861 bp in length  
140454 140503: gap of 50 bp  
140504 143862: contig of 3359 bp in length  
143863 143962: gap of unknown length  
143963 145121: contig of 1159 bp in length  
145122 145518: gap of 397 bp  
145519 147216: contig of 1698 bp in length  
147217 147316: gap of unknown length  
147317 149768: contig of 2452 bp in length  
149769 149818: gap of 50 bp  
149819 152644: contig of 2826 bp in length  
152645 152694: gap of 50 bp  
152695 158870: contig of 6176 bp in length  
158871 159535: gap of 665 bp  
159536 160976: contig of 1441 bp in length  
160977 162113: gap of 1137 bp  
162114 163749: contig of 1636 bp in length  
163750 163849: gap of unknown length  
163850 169431: contig of 5582 bp in length  
169432 169481: gap of 50 bp  
169482 172813: contig of 3332 bp in length  
172814 172913: gap of unknown length  
172914 177020: contig of 4107 bp in length  
177021 177070: gap of 50 bp  
177071 179744: contig of 2674 bp in length  
179745 179912: gap of 168 bp  
179913 182027: contig of 2115 bp in length  
182028 182077: gap of 50 bp  
182078 187508: contig of 5431 bp in length  
187509 187558: gap of 50 bp  
187559 196517: contig of 8959 bp in length  
196518 196567: gap of 50 bp  
196568 200212: contig of 3645 bp in length  
200213 200410: gap of 198 bp

Query Match 2.5%; Score 60.2; DB 12; Length 320940;

Best Local Similarity 59.4%; Pred. No. 7.1e-07;

Matches 142; Conservative 0; Mismatches 88; Indels 9; Gaps 2;

QY 1492 CCTTGACAGATGACGACGATGAGCTTACCGGACGACATCAGT-AAAATGC 1550  
DB 162201 CATTAATGAGTCAAGGCGCTTGAACCTTAACCTTCACACATATGAGCTGG 162260  
QY 1551 TTCTCGATAGCCTGTTCATTAATTCCTTGGCAAGCTCGTACTTACGAGA----- 1605

Db	Accession	Score	Length	Indels	Gaps	Matches	Conservative	Matches	Indels	Gaps
Db	162261	1141 bp	DNA	linear	PAT 14-DEC-2004	162320				
Qy	1606	2.4%; Score 57.8; DB 2; Length 1141;				1662				
Db	162321	10.9%; Pred. No. 4e-06;				162380				
Qy	1633	263; Mismatches 407; Indels 0; Gaps 0;				1721				
Db	162381					162439				
RESULT 10										
AR579680	AR579680	1141 bp	DNA	linear	PAT 14-DEC-2004					
LOCUS	Sequence 22 from patent US 6784342.									
DEFINITION	AR579680									
ACCESSION	AR579680.1	GI:56583130								
VERSION										
KEYWORDS										
SOURCE	Unknown.									
ORGANISM	Unknown.									
REFERENCE	Unclassified.									
AUTHORS	1 (bases 1 to 1141)									
TITLE	Kunet, L. and Clemens, S.									
JOURNAL	Regulation of embryonic transcription in plants									
	Patent: US 6784342-A 22 31-AUG-2004;									
	The University of British Columbia; Vancouver;									
	CAX;									
FEATURES	location/Qualifiers									
source	1..1141									
	/organism="unknown"									
	/mol_type="genomic DNA"									
ORIGIN										
Query Match	2.4%; Score 57.8; DB 2; Length 1141;									
Best Local Similarity	10.9%; Pred. No. 4e-06;									
Matches	82; Conservative 263; Mismatches 407; Indels 0; Gaps 0;									
Qy	1565	TGTTGATTAATTCCTCTTGGCAAAGCTGCTACTTACAGAAAGTCGCTACAGACAGA	1624							
Db	201	DRDMSBRKMYGMBWKMWSYDYVYVWWDMDCKRVRVTRGRMRNTWVMBTJHR	260							
Qy	1625	AGATGCGATTTACGATCTTGAAAAATAGAAAAATAGAACTATGAAAAATTCATGG	1684							
Db	261	RRYNGTBTBAAHYRRWTNNNNNNNAKMKCKAKWGMRAVNSCTTWSKTKYRTSC	320							
Qy	1685	AAGCAATGGGTAAGCTTACTTTTGAATGCTTTCTAAAGCAGATACACCTACGGCG	1744							
Db	321	WANNCRAGDAKDKHKWKMWSAAMGYVNNNNNNNTYKARBARBWDVWMSAMKWHAN	380							
Qy	1745	GAATACAACTTAAGCTGTTCATGAATCAATCGCTGCTAAGCTGCTCTTGTGTGTCG	1804							
Db	381	AAHSRRKKTWYKRTTVNNNNNGTTMKRMAWYMKMDMGTTNNNNNGRTYYGWTK	440							
Qy	1805	TATTTGGCCCTTGACATTCGCCCTCACTTAATTTGAAAAAGCTATAGAGGGAATTC	1864							
Db	441	NKKWWTYYKWKANCKRPMDBKCTTHNNVTTWKKTYNNNCYKWSMTGKSHRBAAY	500							
Qy	1865	AAGGAAGAAAAACATTCGATTTATTTGATTTGCGATGATATCTTAATGATTAAGCTAAT	1924							
Db	501	TWYMMWRRYHANNNNNDYMKACTWYBYCSKMMNNYAAWYTKSSWNTSRYYRMKT	560							
Qy	1925	CCAGTAGAGGCAATTCACAGAAATTTAAATGAATATATGTAAGAAATATATTTTGA	1984							
Db	561	NNSWWRSDTBSMGRANNYARABHHYQKNTFTWMBSHWBHRAAAYWMBMVBAC	620							
Qy	1985	TAAAGCTGTTGAAAAATTAACAACAAGGGAATTCGTGTCGACGTTTGGAGAACAC	2044							
Db	621	HCKMAYAKKYAAGGSSNNNNNNNNNNNNNNNNNNATCAADDYTAASRWYMAAAKYYK	680							
Qy	2045	ACATATTTGAATCTATTTTAACATGCTAGTCTTACTTTAAGCTTCAAACTGCTGT	210							

QY	2105	AAATGATGTAAACATACTTCTAGTGTAGTACTGTTGATTTACAGGCTGAACAC	2164
Db	741	AAVAAKHKRMAMKAMRGMBADAAABFTDDKKNCAAYTKYTTTNNNTYRGVYTTAARD	800
QY	2165	TGCGCTCAGTGAAGGATGAGAGAGTGAAGTCTGAGTCAGAAATTCGGGCTAACCTCCC	2224
Db	801	GMAMNNNNNNNNNNNNNGWSDMWVYTWMAVAYGTNNNNNNNNNNNVAWYTWKMYTTDTR	860
QY	2225	TCAACTACAGAAAAAGTCACAATATAAATATGCAACATGATGTTCTCATATTTGTTTTCTCT	2284
Db	861	WRBAYTNNNNNNNNMAYGAYADYAAWMSDTCDAWAKMDATKNNATYYNRTGTAWRTNNN	920
QY	2285	GCTTGATGTTAATTGATTTATTTATTTATTTT	2316
Db	921	NNNNYTKTYBHAAMNNNNNNCKMCTATHTW	952
RESULT	11		
LOCUS	AX083744	1141 bp	DNA linear PAT 28-FEB-2001
DEFINITION	Sequence 22 from Patent WO0111061.		
ACCESSION	AX083744		
VERSION	AX083744.1	GI:13185472	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS	Kunst, L. and Clemens, S.		
TITLE	Regulation of embryonic transcription in plants		
JOURNAL	Patent: WO 0111061-A 22 15-FEB-2001; UNIVERSITY OF BRITISH COLUMBIA (CA)		
FEATURES			
source	1. 1141 /organism="synthetic construct" /mol_type="unassigned DNA" /db_xref="taxon:32630"		
	1. 1141 /note="consensus sequence of A.t., L.a., and B.n. FAE1 promoters"		
promoter			
	2.4%: Score 57.8; DB 2; Length 1141; Best Local Similarity 10.9%; Pred. No. 4e-06; Matches 82; Conservative 263; Mismatches 407; Indels 0; Gaps 0;		
QY	1565	TGTTCATAAATCTCTTTGCAAGCTCTGCTACTTACACAGAGTCTGCTTACAGACAGAA	1624
Db	201	DRDMSBKMYNGMBWKKMSYDYVYVWVMDCKRVRWRVTRGMRVRYMAWBYAHR	260
QY	1625	AGATGGCATTTAACGCTACTTGAAAAATAGAAAAATGAGAACTATGAAAAATTCATCG	1684
Db	261	RRYNNGMTBAAYRRYTNNNNNNNNAKAMCKAKTKGMRBAYNSTCTTWSKTKVTSC	320
QY	1685	AAGCATGGGTAAAGCTACTTTTGTGAATGCTCTTAAAGACGATPACACATACGGCG	1744
Db	321	WANNCRAGDANKDHKMKMSAMGYVNNNNNNNNMTYKKAHBARMVWMSAKKMHAN	380
QY	1745	GAATCAAACTTAACGTGTATGAGAACATACGCTGCTACCTGCTCTTGTGTGTGC	1804
Db	381	AAHYSRKMTBKRKTMTVNNNNNGTTWIKMAWYKMKMDMBGTNNNNNNNGRYYGWTX	440
QY	1805	TATTTGACCTTGACATTTGGCCCTGACCTTAATTTTGAAGAAGCTCTATAGAGGGAATAC	1864
Db	441	NKKMTYYMKKANKMKRMBHKCTHTNNHTTWMKKTYMNNCYSMSMTNGSHBBAAYV	500
QY	1865	AAGGAAGAAAACATTCGATTTTATTTGACATGGAATPACTTATGATTTAGTATTAAT	1924
Db	501	TWTMMWRRYHAJANNNNMDYWMKACTWYKYBVCISKMNNYAAWYTKSSWNTTSRYRKKT	560
QY	1925	CCAGTGAGAGCAATCCAGAGAAATTTAAATGAATTTATGTAGAATAATTAATTTGA	1984

[illegible]

Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,  
 Sutton, A., Szatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
 Tansley, J., Taylor, C., Taylor, F., Telirod, B., Thomas, N., Thomas, S.,  
 Umanan, K., Vasquez, J., Vera, V., Verdusco, D., Villalon, D.,  
 Vinsan, R., Waldron, L.M., Wang, Q., Wang, S., Ward-Moore, S.,  
 Warren, R., Washington, C., Watlingsson, S., Williams, G.,  
 Williamson, A., Wiczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y.,  
 Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G. and Gibbs, R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 320940)  
 Worley, K.C.  
 Direct Submission  
 Submitted (18-NOV-2004) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 320940)  
 Worley, K.C.  
 Direct Submission  
 Submitted (23-JAN-2005) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jan 23, 2005 this sequence version replaced gi:55819826.  
 The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
 in the feature table below represents a scaffold in the Atlas  
 assembly (a 'contig-scaffold'). Within each contig-scaffold,  
 individual sequence contigs are ordered and oriented, and separate  
 by sized gaps filled with Ns to the estimated size. The sequence  
 may extend beyond the ends of the clone and there may be sequence  
 contigs within a contig-scaffold that consist entirely of whole  
 genome shotgun sequence reads. Both end sequences and whole genome  
 shotgun sequence only contigs will be indicated in the feature  
 table.

---

----- Genome Center -----  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information -----  
 Center project name: FARP  
 Center clone name: CH240-10M22  
 ----- Summary Statistics -----  
 Assembly program: Atlas 3.0;  
 Consensus quality: 288280 bases at least Q40  
 Consensus quality: 294971 bases at least Q30  
 Consensus quality: 301063 bases at least Q20  
 Estimated insert size: 292138; sum-of-contigs estimation  
 Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

---

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html))  
 \* NOTE: This sequence may represent more than one clone.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 81 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1	5541: contig of 5541 bp in length
5542	5591: gap of 50 bp
5592	5604: contig of 4013 bp in length
9605	9704: gap of unknown length
9705	11503: contig of 1799 bp in length
11504	11802: gap of 299 bp
11803	20154: contig of 8352 bp in length
20155	20254: gap of unknown length
20255	24517: contig of 4263 bp in length
24518	24567: gap of 50 bp
24568	25759: contig of 1192 bp in length

```

* 25760 26270: gap of 511 bp
* 26271 29359: contig of 3089 bp in length
* 29360 29459: gap of unknown length
* 29460 32345: contig of 2886 bp in length
* 32346 32395: gap of 50 bp
* 32396 38577: contig of 6182 bp in length
* 38578 38678 40965: contig of 2288 bp in length
* 40966 41015: gap of 50 bp
* 41016 47383: contig of 6368 bp in length
* 47384 47484 48761: gap of unknown length
* 48762 48842: gap of 81 bp
* 48843 51170: contig of 2328 bp in length
* 51171 51270: gap of unknown length
* 51271 51292: contig of 1822 bp in length
* 53093 53142: gap of 50 bp
* 53143 55224: contig of 2082 bp in length
* 55225 55313: gap of 89 bp
* 55314 57275: contig of 1962 bp in length
* 57276 57375: gap of unknown length
* 57376 63877: contig of 6502 bp in length
* 63878 63927: gap of 50 bp
* 63928 65212: contig of 1285 bp in length
* 65213 65262: gap of 50 bp
* 65263 66517: contig of 1255 bp in length
* 66518 66617: gap of unknown length
* 66618 70198: contig of 3581 bp in length
* 70199 71339: gap of 1141 bp
* 71340 75208: contig of 3869 bp in length
* 75209 75309 97282: gap of unknown length
* 75309 97282: contig of 21924 bp in length
* 97283 97283: gap of 50 bp
* 97283 100331: contig of 3049 bp in length
* 100332 101076: gap of 745 bp
* 101077 102495: contig of 1419 bp in length
* 102496 102595: gap of unknown length
* 102596 104484: contig of 1889 bp in length
* 104485 104708: gap of 224 bp
* 104709 109824: contig of 5116 bp in length
* 109825 109924: gap of unknown length
* 109925 115197: contig of 5273 bp in length
* 115198 115247: gap of 50 bp
* 115248 123071: contig of 7824 bp in length
* 123072 123171: gap of unknown length
* 123172 131072: contig of 7901 bp in length
* 131073 131122: gap of 50 bp
* 131123 135372: contig of 4250 bp in length
* 135373 135472: gap of unknown length
* 135473 140453: contig of 4981 bp in length
* 140454 140503: gap of 50 bp
* 140504 143862: contig of 3359 bp in length
* 143863 143962: gap of unknown length
* 143963 145121: contig of 1159 bp in length
* 145122 145216: gap of 397 bp
* 145217 147216: contig of 1698 bp in length
* 147217 147316: gap of unknown length
* 147317 149768: contig of 2452 bp in length
* 149769 149818: gap of 50 bp
* 149819 152644: contig of 2826 bp in length
* 152645 152694: gap of 50 bp
* 152695 158870: contig of 6176 bp in length
* 158871 159535: gap of 665 bp
* 159536 160976: contig of 1441 bp in length
* 160977 162113: gap of 1137 bp
* 162114 163749: contig of 1636 bp in length
* 163750 163849: gap of unknown length
* 163850 169431: contig of 5582 bp in length
* 169432 169481: gap of 50 bp
* 169482 172813: contig of 3332 bp in length
* 172814 172913: gap of unknown length
* 172914 177020: contig of 4107 bp in length
* 177021 177070: gap of 50 bp

```

Query Match 2.4%; Score 57.6; DB 12; Length 320940;  
 Best Local Similarity 65.6%; Pred. No. 4.7e-06;  
 Matches 84; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

```

QY 1594 CTACTTACGAGAAAGTCTGCTTACAGACGAAAGATGAGCTTTACGTAAGTGA 1653
    |||
Db 22649 CTGATTGCTCAGAGGGCTGAATTACACGACACCATGGCTTTAAAGTACTGGAAGTTA 22590
    |||
QY 1654 GAGAAATGAGAACTATGAAAATTCATGAGAAAGCAATGGTAAAGCTTACTTTTGA 1713
    |||
Db 22589 GACAGAAATGAGAACTATGAAAAGTTCATGGGAAAAGGTAAGTGTAACTACTTAA 22530
    |||
QY 1714 TGCCTTCT 1721
    |||
Db 22529 CAGCTTGT 22522
    |||

```

## RESULT 13

MUSFABPI  
 LOCUS MusFABPI gene, exons 1-4.  
 DEFINITION Mouse FabpI  
 ACCESSION M65033.1 GI:193218  
 VERSION M65033.1  
 KEYWORDS Mus musculus (house mouse)

ORGANISM Mus musculus (house mouse)  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Muroidae; Muridae; Murinae; Mus.

## REFERENCE

1 (sites)  
 Green, R.P., Cohn, S.M., Sacchetti, J.C., Jackson, K.E. and  
 Gordon, J.I.  
 The mouse intestinal fatty acid binding protein gene: nucleotide  
 sequence, pattern of developmental and regional expression, and  
 proposed structure of its protein product  
 DNA Cell Biol. 11 (1), 31-41 (1992)

## JOURNAL

PUBMED

1739433

## REFERENCE

AUTHORS

2 (bases 1 to 5039)  
 Green, R.P., Sacchetti, J.C., Jackson, K.E., Cohn, S.M. and  
 Gordon, J.I.

## TITLE

The mouse fatty acid binding protein gene (FabpI): nucleotide  
 sequence analysis and predictions of the structure of its protein  
 product

## JOURNAL

Unpublished (1991)

## COMMENT

Original source text: Mus musculus (strain DBA/2J) Adult Liver DNA.

## FEATURES

source

Location/Qualifiers

1..5039

FEATURES

/organism="Mus musculus"

FEATURES

/mol\_type="genomic DNA"

FEATURES

/strain="DBA/2J"

FEATURES

/db\_xref="taxon:10090"

FEATURES

/tissue\_type="Liver"

FEATURES

/dev\_stage="Adult"

FEATURES

467..547

FEATURES

/note="Domain I: homology with human I-FABP promoter  
 region"; putative"

FEATURES

524..537

FEATURES

/note="one of two repeats homologous to consensus repeat  
 found in number of other lipid binding proteins which are  
 expressed in small intestine"; putative"

FEATURES

727..817

FEATURES

/note="Domain II: homology to human I-FABP promoter"

```

misc_feature      region: putative"
                   870..966
                   /note="Domain III: homology with human I-FABP promoter
                   region"; putative"
repeat_region     /note="Second of two repeats homologous to consensus
                   repeat; putative"
                   884..897
                   /note="putative"
                   937..942
TATA_signal       /note="putative"
                   967..1069
                   /label=Exon1
exon              /label=Exon1
                   2326..2497
                   /note="putative"
                   /label=Exon2
exon              /note="putative"
                   4291..4398
                   /label=Exon3
exon              /label=Exon3
                   4818..5039
                   /note="putative"
                   /label=Exon4

ORIGIN
Query Match      2.4%; Score 56.6; DB 6; Length 5039;
Best Local Similarity 58.5%; Pred. No. 9.6e-06;
Matches 121; Conservative 0; Mismatches 79; Indels 7; Gaps 1;

QY 1516 TGAGCTTAGCAGCAGCATATCATATGTAATTGCTTCCTGATAGAGCTGTTCATATAT 1575
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 884 TGAACCTTGCAACTTCCACATCATCAAGTATGAAATGGTTGGAAGATGAAGAAATGAATTAAT 943
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 1576 TCTCTTTGCAAGCTCTGCTACTTACCAAGATC-----TGCTTACAGACAGAGAAAGAT 1628
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 944 TCTCTTAGAGTGAAGAGAGCTGAGACTGCTTCTTCTTAGAGACAGACAGCTGAGATCAT 1003
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 1629 GGCATTTAACGCTACTCTTGGAAAAATAGAGAAAAATGAGACTATGAAAAAATTCATGGAAC 1688
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1004 GGCATTTCGACGCGACGTGGAAGAGAGACCGGAGACGAGAACTATGAAAAAGTTTCATGAGAA 1063
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 1689 AATGGTAGCGCTTACTTTTGAATG 1715
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1064 AATGGTAGCGCTTACTTCTGCTG 1090
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 14
AC158770/c      188707 bp DNA linear HTG 23-MAR-2005
LOCUS           Mus musculus chromosome 1 clone RP23-393018 map 1, WORKING DRAFT
DEFINITION     AC158770
SEQUENCE       AC158770.1 GI:61696497
VERSION        HTG; HTGS PHASE1; HTGS_DRAFT.
KEYWORDS       Mus musculus (house mouse)
SOURCE         Mus musculus
ORGANISM       Mus musculus
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
TITLE          Sciurognathi; Muroidae; Muridae; Murinae; Mus.
JOURNAL        1 (bases 1 to 188707)
REFERENCE      2 (bases 1 to 188707)
AUTHORS        Birren, B., Nusbaum, C. and Lander, E.
TITLE          Mus musculus chromosome 1, clone RP23-393018
JOURNAL        Unpublished
AUTHORS        2 (bases 1 to 188707)
REFERENCE      Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
AUTHORS        Anderson, M., Anderson, S., Arachchi, H. M., Barina, N., Bastien, V.,
TITLE          Bloom, T., Boguslavsky, L., Bouhagalter, B., Camarata, J., Chang, J.,
JOURNAL        Choepel, T., Collamore, A., Cook, A., Cooke, P., Corum, B.,
AUTHORS        Dearellano, K., Diaz, J. S., Dodge, S., Doolley, K., Dorris, L.,
TITLE          Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D.,
JOURNAL        Galagan, J., Gardner, S., Graham, L., Grand-Pierre, N., Hafez, N.,
AUTHORS        Hagopian, D., Hages, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,
TITLE          Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, T.,
JOURNAL        Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
AUTHORS        Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C.,

```

```

TITLE
JOURNAL
COMMENT
-----
Center: Broad Institute of MIT and Harvard
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@broad.mit.edu
-----
Project Information
Center project name: 131651
Center clone name: 393_O_18
-----
Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 183196 bases at least Q40
Consensus quality: 185063 bases at least Q30
Consensus quality: 186541 bases at least Q20
Insert size: 230000; agarose-fp
Insert size: 187207; sum-of-contigs
Quality coverage: 4.5 in Q20 bases; agarose-fp
Quality coverage: 5.5 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
-----
1
52662: contig of 52662 bp in length
52663: gap of unknown length
52763: contig of 100 bp in length
52862: contig of 100 bp in length
52863: gap of unknown length
52963: contig of 954 bp in length
53917: gap of unknown length
54016: gap of unknown length
54017: contig of 2079 bp in length
56095: gap of unknown length
56195: gap of unknown length
56196: contig of 1697 bp in length
57892: contig of 1697 bp in length
57893: gap of unknown length
61597: contig of 3605 bp in length
61598: gap of unknown length
61599: gap of unknown length
64421: contig of 4224 bp in length
64422: gap of unknown length
66522: gap of unknown length
66523: contig of 4268 bp in length
70789: gap of unknown length
70790: gap of unknown length
74949: contig of 4060 bp in length
74950: gap of unknown length
74951: gap of unknown length
75050: gap of unknown length
75051: contig of 6908 bp in length
81958: gap of unknown length
82057: gap of unknown length
82058: contig of 7821 bp in length
89879: gap of unknown length
89879: gap of unknown length
103099: contig of 13121 bp in length
103100: gap of unknown length
103101: contig of 10948 bp in length
114247: gap of unknown length
114248: gap of unknown length
114248: contig of 16591 bp in length
130838: gap of unknown length
130839: gap of unknown length
130938: gap of unknown length
162497: contig of 31559 bp in length

```

TITLE  
 JOURNAL  
 COMMENT  
 Submitted (23-MAR-2005) Broad Institute of MIT and Harvard, 320  
 Charles Street, Cambridge, MA 02141, USA  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 -----

Center: Broad Institute of MIT and Harvard  
 Center code: MIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence.submissions@broad.mit.edu  
 -----

Project Information  
 Center project name: 131651  
 Center clone name: 393\_O\_18  
 -----

Summary Statistics  
 Sequencing vector: Plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 183196 bases at least Q40  
 Consensus quality: 185063 bases at least Q30  
 Consensus quality: 186541 bases at least Q20  
 Insert size: 230000; agarose-fp  
 Insert size: 187207; sum-of-contigs  
 Quality coverage: 4.5 in Q20 bases; agarose-fp  
 Quality coverage: 5.5 in Q20 bases; sum-of-contigs  
 -----

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 16 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1  
 52662: contig of 52662 bp in length  
 52663: gap of unknown length  
 52763: contig of 100 bp in length  
 52862: contig of 100 bp in length  
 52863: gap of unknown length  
 52963: contig of 954 bp in length  
 53917: gap of unknown length  
 54016: gap of unknown length  
 54017: contig of 2079 bp in length  
 56095: gap of unknown length  
 56195: gap of unknown length  
 56196: contig of 1697 bp in length  
 57892: contig of 1697 bp in length  
 57893: gap of unknown length  
 61597: contig of 3605 bp in length  
 61598: gap of unknown length  
 61599: gap of unknown length  
 64421: contig of 4224 bp in length  
 64422: gap of unknown length  
 66522: gap of unknown length  
 66523: contig of 4268 bp in length  
 70789: gap of unknown length  
 70790: gap of unknown length  
 74949: contig of 4060 bp in length  
 74950: gap of unknown length  
 74951: gap of unknown length  
 75050: gap of unknown length  
 75051: contig of 6908 bp in length  
 81958: gap of unknown length  
 82057: gap of unknown length  
 82058: contig of 7821 bp in length  
 89879: gap of unknown length  
 89879: gap of unknown length  
 103099: contig of 13121 bp in length  
 103100: gap of unknown length  
 103101: contig of 10948 bp in length  
 114247: gap of unknown length  
 114248: gap of unknown length  
 114248: contig of 16591 bp in length  
 130838: gap of unknown length  
 130839: gap of unknown length  
 162497: contig of 31559 bp in length

\* 162498 162597: gap of unknown length  
\* 162598 188707: contig of 26110 bp in length.  
Location/Qualifiers

FEATURES  
source  
1.188707

/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="1"  
/map="1"

/clone="RP23-393018"  
/clone\_1lb="RP23-23 Female Mouse BAC"

misc\_feature

1. .52662  
/note="assembly\_fragment"

gap

/clone\_end:SP6  
vector\_side:left"

misc\_feature

52663..52762  
/estimated\_length=unknown

gap

52763..52862  
/note="assembly\_fragment"

misc\_feature

52863..52962  
/estimated\_length=unknown

gap

52963..53916  
/note="assembly\_fragment"

misc\_feature

53917..54016  
/estimated\_length=unknown

gap

54017..56095  
/note="assembly\_fragment"

misc\_feature

56096..56195  
/estimated\_length=unknown

gap

56196..57892  
/note="assembly\_fragment"

misc\_feature

57893..57992  
/estimated\_length=unknown

gap

57993..61597  
/note="assembly\_fragment"

misc\_feature

61598..61697  
/estimated\_length=unknown

gap

61698..66421  
/note="assembly\_fragment"

misc\_feature

66422..66521  
/estimated\_length=unknown

gap

66522..70789  
/note="assembly\_fragment"

misc\_feature

70790..70889  
/estimated\_length=unknown

gap

70890..74949  
/note="assembly\_fragment"

misc\_feature

74950..75049  
/estimated\_length=unknown

gap

75050..81957  
/note="assembly\_fragment"

misc\_feature

81958..82057  
/estimated\_length=unknown

gap

82058..89878  
/note="assembly\_fragment"

misc\_feature

89879..89978  
/estimated\_length=unknown

gap

89979..103099  
/note="assembly\_fragment"

misc\_feature

103100..103199  
/estimated\_length=unknown

gap

103200..114147  
/note="assembly\_fragment"

misc\_feature

114148..114247  
/estimated\_length=unknown

gap

114248..130838  
/note="assembly\_fragment"

misc\_feature

130839..130938  
/estimated\_length=unknown

gap

130939..162497  
/note="assembly\_fragment"

misc\_feature

162498..162597  
/estimated\_length=unknown

misc\_feature 162598..188707  
/note="assembly\_fragment"  
clone\_end:77  
vector\_side:right"

ORIGIN

Query Match 2.4%; Score 56.6; DB 12; Length 188707;  
Best local Similarity 58.5%; Pred. No. 9.7e-06;  
Matches 121; Conservative 0; Mismatches 79; Indels 7; Gaps 1;

QY

1516 TGAGCTTTAGCCAGCAGCATCATCATGTAATTTGCTTCCGATATAGCCTGTTCAAT 1575

Db

83524 TGAACCTTCAGACTTCACATTCACAGTATGATTTGTTGCAAGATTAAGAAATGAAAT 83465

QY

1576 TCTCTTGGAAAGCTCTGCTACTTACCGAAGTC-----TGCTTACGACGAAAGAT 1628

Db

83464 TCTCTTACGTGACAGAGACTGACCTCTCTTTCCTAGAGACACACAGCTGAGATCAT 83405

QY

1629 GGCATTAAAGCTACTTGGAAATAGAAAAATGAAACTATGAAAAATTCATGGAAC 1688

Db

83404 GGCATTGACGCGACGTGAAAGTACCGGAACGAACTATGAAAGTTTCATGGAAGA 83345

QY

1689 AATGGTAAGCCTTACTTTTGAATG 1715

Db

83344 AATGGTAAGCACTGCTCTCTCTG 83318

RESULT 15

AC155158/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REVIEW

AUTHORS

TITLE

JOURNAL

REVIEW

AUTHORS

TITLE

JOURNAL

REVIEW

AUTHORS

TITLE

JOURNAL

REVIEW

AUTHORS

TITLE

JOURNAL

REVIEW

AUTHORS

TITLE

JOURNAL

REVIEW

AUTHORS

TITLE

JOURNAL

REVIEW

AUTHORS

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Submitted (22-Oct-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 240767)  
Birken, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, N., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
COMMENT

Submitted (10-Nov-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA  
5 (bases 1 to 240767)  
Birken, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

FEATURES  
source  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Broad Institute of MIT and Harvard  
Center code: WMRB  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@broad.mit.edu  
----- Project Information  
Center project name: U30890  
Center clone name: 70\_L\_3  
-----  
Location/Qualifiers  
1. 240767  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="3"  
/map="3"  
/clone="RP24-70L3"  
/clone\_id="RPC1-24 Male Mouse BAC"  
110. 298  
/rpt\_family="L2"  
complement(11767..2017)  
/rpt\_family="B4A"  
2076..2104  
/rpt\_family="AT\_rich"  
2157..2205  
/rpt\_family="AT\_rich"  
2305..2398  
/rpt\_family="Lx4"  
complement(3241..3329)  
/rpt\_family="L1MEC"  
3371..3457  
/rpt\_family="PB1D10"  
3699..3738  
/rpt\_family="T-rich"  
3825..3942  
/rpt\_family="B1\_MM"  
complement(4160..4441)  
/rpt\_family="L1MEC"  
complement(4648..4730)  
/rpt\_family="MIR"  
complement(4768..4850)  
/rpt\_family="MIR"  
5595..5637  
/rpt\_family="T-rich"  
complement(6332..6443)  
/rpt\_family="L2"  
complement(6768..7027)  
/rpt\_family="B4A"  
7531..7581  
/rpt\_family="AT\_rich"  
complement(7649..7898)  
/rpt\_family="Lx8"  
8014..8135  
/rpt\_family="B4"  
8197..8370  
/rpt\_family="B2\_MM2"  
8725..8753  
/rpt\_family="(T)n"  
10512..10618  
/rpt\_family="PB1D7"  
10654..10845  
/rpt\_family="RMER17C"  
11363..11475  
/rpt\_family="PB1"  
complement(11517..11546)  
/rpt\_family="B3"  
complement(11547..11695)  
/rpt\_family="B1\_MM"  
complement(11696..11877)  
/rpt\_family="B3"  
complement(11894..11987)  
/rpt\_family="B1F"

```
repeat_region complement(13781, .13925)
/rpt_family="B1F"
repeat_region 13983, .14157
/rpt_family="(TG)n"
repeat_region 14188, .14320
/rpt_family="L1ME1"
unsure 14247
/note="single clone coverage"
repeat_region 14615, .14642
/rpt_family="(A)n"
```

Query Match 2.4%; Score 56.6; DB 6; Length 240767;

Best Local Similarity 58.5%; Pred. No. 9.8e-06; Mismatches 79; Indels 7; Gaps 1;

Matches 121; Conservative 0;

```
QY 1516 TGAGCTTTAGCCGACATCATGTAATTTGCTTCTCGATTAAGCTGTTCTATTAAT 1575
DB 74906 TGAACCTTCGAACTTCACATCAGTATGTAATTGTTTGAAGATTAAGAAATTAAT 74847
QY 1576 TCTCTTGCAGGCTCTGCTACTTACCGAAGTC-----TGCTTACAGACAGAAAGAT 1628
DB 74846 TCTCTTACAGGACAGACGACCTCTGCTTCTTCTAGAGACACACAGCTGAGATCAT 74787
QY 1629 GGCATTTAAGGCTACTTGAAGAAATAGAGAAATAGAACTATGAAATTCATGGAAGC 1688
DB 74786 GGCATTCGACGGCAGTGAAGAAAGTAGACCGGACGAGAACTATGAAAGTTTCAATGAGAA 74727
QY 1689 AATGGTAAAGCCTTACTTTTGAATG 1715
DB 74726 AATGGTAAAGCCTGCTTCTGCTG 74700
```

Search completed: May 28, 2006, 10:39:22

Job time : 13882.8 secs





CC gut-specific gene expression control region operably linked to a nucleic  
CC insert encoding a polypeptide, an expression vector that integrates  
CC into a host cell (and comprising the isolated avian gut-specific gene  
CC expression control region), expressing a heterologous polypeptide in a  
CC host cell (by transfecting a eukaryotic cell with the recombinant DNA  
CC molecule, and culturing the transfected cell in a medium suitable for  
CC expression of a heterologous polypeptide under the control of an avian  
CC intestinal fatty acid binding protein (IFABP) or cp35 gene expression  
CC control region encoded by the recombinant DNA molecule), a eukaryotic  
CC cell transformed with the expression vector (or its progeny, which  
CC expresses a heterologous polypeptide) and a transgenic avian having a  
CC heterologous polynucleotide sequence comprising the nucleic acid insert.  
CC The nucleic acids are useful for regulating heterologous nucleic acids in  
CC transgenic avians, as probes in nucleic acid hybridisation assays for  
CC detecting the IFABP gene expression control region, and for generating  
CC transgenic birds. The present sequence is the chicken intestinal fatty  
CC acid binding protein, IFABP, gene, 5' region.  
XX

SQ Sequence 2381 BP; 775 A; 428 C; 469 G; 709 T; 0 U; 0 Other;

Query Match 100.0%; Score 2381; DB 11; Length 2381;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 2381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCTTCGCGCAGAAAAGGCTGTGGGGTCTTGTTCCTCAGACAGCTTAAGCAAAATC 60  
DB 1 AGCTTCCTCGCAGAAAAGGCTGTGGGGTCTTGTTCCTCAGACAGCTTAAGCAAAATC 60  
QY 61 CCCAAGTTCAAAACGTGCGCTGTAAAGAGAGATGCTCACTTCAAAATGAAGTAATG 120  
DB 61 CCCAAGTTCAAAACGTGCGCTGTAAAGAGAGATGCTCACTTCAAAATGAAGTAATG 120  
QY 121 AAATATCTAATAACGAGCTCTGTGGCAGATCAGAGATTAACCTTGTGGGAGCAAAAT 180  
DB 121 AAATATCTAATAACGAGCTCTGTGGCAGATCAGAGATTAACCTTGTGGGAGCAAAAT 180  
QY 181 CTTAAGGTATAGGTAGAGACAGAGGTGTTCGAACTTAATGAGACTTGAATGACACA 240  
DB 181 CTTAAGGTATAGGTAGAGACAGAGGTGTTCGAACTTAATGAGACTTGAATGACACA 240  
QY 241 TTTGATCTTCTAGAGACAAAAGGCTCTGGAACCAATTAATCTGGTGACAGTCACT 300  
DB 241 TTTGATCTTCTAGAGACAAAAGGCTCTGGAACCAATTAATCTGGTGACAGTCACT 300  
QY 301 AGCAGCCTGTTGGGTGCACTACAGACCTTGTGGCAACATTAATCAATCTAAGTTG 360  
DB 301 AGCAGCCTGTTGGGTGCACTACAGACCTTGTGGCAACATTAATCAATCTAAGTTG 360  
QY 361 TTTTCTTTCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 420  
DB 361 TTTTCTTTCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 420  
QY 421 TACTTCCCTGCACTCTCTTACAGCAATTAAGCAATTAATTAATTAATTAATTAATTA 480  
DB 421 TACTTCCCTGCACTCTCTTACAGCAATTAAGCAATTAATTAATTAATTAATTAATTA 480  
QY 481 ATGGAAGCCACACTATTTTGGTATGACAGCAAGTTCTAATTTATAGTTATTAATCTCA 540  
DB 481 ATGGAAGCCACACTATTTTGGTATGACAGCAAGTTCTAATTTATAGTTATTAATCTCA 540  
QY 541 GTAAACCTTTTGTGCGAGTCTGGAAGAAAAGAAATATATTAACACAGACTAGTA 600  
DB 541 GTAAACCTTTTGTGCGAGTCTGGAAGAAAAGAAATATATTAACACAGACTAGTA 600  
QY 601 AAATTCATAGTATGAGCCAAACCTGTATCTGTGATTAAGCAATTCATTTTCAGCAAT 660  
DB 601 AAATTCATAGTATGAGCCAAACCTGTATCTGTGATTAAGCAATTCATTTTCAGCAAT 660  
QY 661 TCAAGATTTACATTTTGAAGTAAATAGACAGAGATTTGGTCCGTCAATAGAACAGA 720  
DB 661 TCAAGATTTACATTTTGAAGTAAATAGACAGAGATTTGGTCCGTCAATAGAACAGA 720  
QY 721 CTAACTATATCTGAGTTTAGTACAAGAGATTTAGCACAGCAAAATTTGCTCAGTTTC 780

DB 721 CTAACATTAATCTGAGTTTAGTACAAGAGATTTAGCACAGCAAAATTTGCTCAGTTTC 780  
QY 781 AAGTAGCACTATCTTGGGGGAGAAAGAGACTGAGCCAGTGTGCTCATTTTCTCAAT 840  
DB 781 AAGTAGCACTATCTTGGGGGAGAAAGAGACTGAGCCAGTGTGCTCATTTTCTCAAT 840  
QY 841 ATCTTCAACATTTTAAACCTGGGATCTATGAAATCAAAACAGTGGGTAATAATCACT 900  
DB 841 ATCTTCAACATTTTAAACCTGGGATCTATGAAATCAAAACAGTGGGTAATAATCACT 900  
QY 901 TAGAGACATCACTACTGTAGAAATGACAGAAAACAGACATTCATGAAATGGCTAT 960  
DB 901 TAGAGACATCACTACTGTAGAAATGACAGAAAACAGACATTCATGAAATGGCTAT 960  
QY 961 AATATAGAAATAGTAGAAGGTGCTGAATTTTGAATCTAATCTAATTAAGATGAGACA 1020  
DB 961 AATATAGAAATAGTAGAAGGTGCTGAATTTTGAATCTAATCTAATTAAGATGAGACA 1020  
QY 1021 CGAATGAGAAATATCATCCCAATTTCTGTAGCTCAGCACTAGACTGAAAGTTTCTGAAA 1080  
DB 1021 CGAATGAGAAATATCATCCCAATTTCTGTAGCTCAGCACTAGACTGAAAGTTTCTGAAA 1080  
QY 1081 CTGAACCGAGTTTCCCAACCTACCTGTGAGATGTTCAAGATCTTCAATCTCATGCTTAT 1140  
DB 1081 CTGAACCGAGTTTCCCAACCTACCTGTGAGATGTTCAAGATCTTCAATCTCATGCTTAT 1140  
QY 1141 TATGTGAGTAGAATATGATTTCTCACCAATTAAGATGACAAAGAGATTTGTGTTTT 1200  
DB 1141 TATGTGAGTAGAATATGATTTCTCACCAATTAAGATGACAAAGAGATTTGTGTTTT 1200  
QY 1201 ATCTGTGGGTAAATACGTTTTCTCAGTGTATTAACCCCTCCACAGATTAAGTTC 1260  
DB 1201 ATCTGTGGGTAAATACGTTTTCTCAGTGTATTAACCCCTCCACAGATTAAGTTC 1260  
QY 1261 CTATGCAACAAAGAAATGTCATATCATTTCTTATGTCATTAATTTTCAATAGATA 1320  
DB 1261 CTATGCAACAAAGAAATGTCATATCATTTCTTATGTCATTAATTTTCAATAGATA 1320  
QY 1321 GCCGGTTTTTCTACCACTCAAAATAGATGACAAATGAAATGGTGAAGTACTGTTTA 1380  
DB 1321 GCCGGTTTTTCTACCACTCAAAATAGATGACAAATGAAATGGTGAAGTACTGTTTA 1380  
QY 1381 TAAAGAGAGTAATTAAGATATCATCATTTTGAAGCAATTAAGAGGAGAGATTCAG 1440  
DB 1381 TAAAGAGAGTAATTAAGATATCATCATTTTGAAGCAATTAAGAGGAGAGATTCAG 1440  
QY 1441 CAAACAGTGTCTTAAGAGTGAACAAAGTTAACTAAAGTCAACCCCTCTTGAACA 1500  
DB 1441 CAAACAGTGTCTTAAGAGTGAACAAAGTTAACTAAAGTCAACCCCTCTTGAACA 1500  
QY 1501 GATCAATGCCACAGTGGCTTTAGCCAGCAATCATATGTAATTTGCTTCGTGATA 1560  
DB 1501 GATCAATGCCACAGTGGCTTTAGCCAGCAATCATATGTAATTTGCTTCGTGATA 1560  
QY 1561 AGCTGTTCATTAATCTCTTTGCAAAAGCTCTGTAATTCACAAAGTCTGCTTACAGAC 1620  
DB 1561 AGCTGTTCATTAATCTCTTTGCAAAAGCTCTGTAATTCACAAAGTCTGCTTACAGAC 1620  
QY 1621 AGAAGAGTGAATTAACGCTTAAGTGAATAATGAGAAAATGAGAACTATGAAAAATTC 1680  
DB 1621 AGAAGAGTGAATTAACGCTTAAGTGAATAATGAGAAAATGAGAACTATGAAAAATTC 1680  
QY 1681 ATGGAACAAATGGTAAAGCTTACTTTTGAATGCTTCTTAAGACAGATTCAGACTAC 1740  
DB 1681 ATGGAACAAATGGTAAAGCTTACTTTTGAATGCTTCTTAAGACAGATTCAGACTAC 1740  
QY 1741 GCGGGAATCAAACTTAAGTGTTCATGAACATCTGAGCTAATCTGCTTGTGTGT 1800  
DB 1741 GCGGGAATCAAACTTAAGTGTTCATGAACATCTGAGCTAATCTGCTTGTGTGT 1800  
QY 1801 CTGCTATTTTGGCCTTGACATGCTGCTGACTTAATTTTGAAGAAAGCTTATAGAGGGA 1860





Query Match	2.1*	Score 50.6;	DB 5;	Length 4590;
Best Local Similarity	11.8*	Fred, NO. 0.19,		
Matches 113;	Conservative 416;	Mismatches 423;	Indels 3;	Gaps 3

OY	153	CAGAGATPAACTCGCGTGGGACAAATATTTCTAAAGTPTATAGTGAAGAACAAGAGCTGTTT	212
Db	3544	CCRRNGTCAMSBSTANTPAAYASHRDSRCBBDTHRNCTTAMASANDDRANMGSDATDHS	3603
OY	213	GCAACTAAATGAGCTAGATGTGTACCACTTGTGATCTTCTAGAGACAAAAGGCTGGAA	272
Db	3604	NNTDADADAYSTYMRGRMATTHDAVNTTNTNTANDYAGCTBRNNGRAATVMBRN	3663
OY	273	AACAAATTAATTCGTGTGCACAGTCACTAGTCAAGCCTGTGTTGGTGCATACAGCAACT	332
Db	3664	TRACANTSNANNCASNRGRMATYDASBARCHANDVWNTCTDAND -RAMTNVSTYATH	3722
OY	333	TTGTTTGCAACAATACATCTAAGTGTGTTCTTTCTCTCTCTTAACTTCTGTACA	392
Db	3723	RTYADRSARCHANDNDSTRADVWNTDVBVTHPATNTATNTNYSGRTHNACKMSNSTTTA	3782
OY	393	GTCRAAAGGTGAAGAGTAGCTATTGAGTTACTTCCCTGTGATCCTCTTGAGCCAGATTAG	452
Db	3783	TTNTANDTRADMAKRAATNTNTNTSASTRAAAGAGAGAAAMNRGMAATMBSHMNDYSKDCITMD	3842
OY	453	CATTGATTTTCAAAATGAACTGAGTGAATGAAAGCAACATATTTTGATATCACAG	512
Db	3843	MMYYCKDAKCDOMTNGYCHNNNRBDHDDNDNRVMBRVHDCVWBNVWVWDHDDHNDH	3902
OY	513	-AAAGTTCTAATTTATGTTATGTTATCTTCAGTAAACCTTTTGCTGCAAGTCTGACAA	571
Db	3903	DHDHNDCHNNNBVHDDHDDDDVYNNRGMHNNNNHNTCHDHDHNTCHADGK	3962
OY	572	AAGAAGTTTATGATTAACACAGACTATATAAATTCATTATTTGAGCCAAACCCGTATAC	631
Db	3963	MVVVVVVVVVVVVVVCHNHDHNDHGNHCNCCCKNKHDAANRMAKMNHNHNDG	4022
OY	632	TGTGTGATTAAGCAACTTTCATTTGACATTCAGAGATTGATTTTGAACATAATACAGA	691
Db	4023	CKADATAPARAGRANTTADRTTRAGMBNNNRBSDDMKMMMKMYKNDVWRYT	4082
OY	692	GCAGATTTGTCGCCGTCCATAGAAACAGACTAACTATATCTGAGTTTGTAGTACAGACAG	751
Db	4083	GBDRKNMNDYSMBSHMNDYSKDCMNDYCTMATRCVRSAYASVDNDYSCMTATRCVR	4142
OY	752	ATTTAGACACCAAAATTTGCTCAGTTTCAAGTACACTATCTTGTTGGGGAAGAAGAC	811
Db	4143	YSASVADMDYSCMTATRCVRSAYASVD -WNDYSCMTATRCVRSAYASVDMDYSCMTAT	4201
OY	812	TGACCCAGTGTGTCATTTTCTGCAATTAATCTTCAACATTTTAAACCTGGGATCTATG	871
Db	4202	KCVRYSAVASDMDNDYCTMATRCVRSAYASVDKSKKMSKMSMGMSMWRBANSYTB	4261
OY	872	GAATTCACAACAGTTGGTGAATAATCTACTTAGACAGACATCAACTACTGTAGGAATGAC	931
Db	4262	RACRBNHNBBSDBMSSTDMDRDCSTMATSSCDTWNNDYEMNDYSHHSWNDYSNDNDSC	4321
OY	932	AGAAACAGAGACTTCACTGAATGGGCTATATATAGAAATACGTAGAGGTCTCGAA	991
Db	4322	WNDYSDMCRSTWRBRSDGRTTHNACKHTTDCGABNDDBDNTNRYRSHARSBTABRAC	4381
OY	992	TTTAGACTACTTATTAAGAAGTGAAGACAGAAATGAGAAATATCATGCAATTTCTGTAG	1051
Db	4382	ATMSADMTADMSNTDMDNDSTMDRDCMNTCMNTSKDNTYCDTMTANASADGSTSMARY	4441
OY	1052	CTCAGCACTAGACTCGAAGGTTTCTGAATCTGAACGAGCTTCCAAACTACTCTG	1106
Db	4442	NBMATNNNSTRTATTCANABRRATWDCMNTSMABRYRMATNRSCHMTCDACMBS	4496
RESULT 5			
ABK83946			
ID ABK83946 standard; cDNA, 5204 BP.			
XX ABK83946;			
XX			
XX 14-AUG-2002 (first entry)			
XX			

DE Human cDNA differentially expressed in granulocytic cells #517.  
XX  
XX Human; 89; granulocytic cell; DNA chip; bacterial infection;  
KM viral infection; parasitic infection; protozoal infection;  
KM fungal infection; sterile inflammatory disease; psoriasis;  
KM rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;  
KM cardiac reperfusion injury; renal reperfusion injury; ARDS;  
KM adult respiratory distress syndrome; inflammatory bowel disease;  
KM Crohn's disease; ulcerative colitis; periodontal disease;  
KM granulocyte activation; chronic inflammation; allergy.  
XX  
OS Homo sapiens.  
XX  
PN WO200228999-A2.  
XX  
PD 11-Apr-2002.  
XX  
PF 03-OCT-2001; 2001MO-US030821.  
XX  
PR 03-OCT-2000; 2000US-0237189P.  
XX  
PA (GENE-) GENE LOGIC INC.  
XX  
PI Beazer-Barclay Y, Weisman SM, Yamaga S, Vockley J;  
XX WPI; 2002-435328/46.  
XX  
PT Detecting granulocyte activation by detecting differential expression of  
PT genes associated with granulocyte activation, which serves as diagnostic  
PT markers that is useful for monitoring disease states and drug toxicity.  
XX  
PS Claim 1, SEQ ID NO 517; 114pp; English.  
XX  
XX The invention relates to detecting (M1) granulocyte (GC) activation  
XX (GCA), by detecting the level of expression of gene(s) (Gs) identified by  
XX DNA chip analysis as given in the specification, and comparing the  
XX expression level to an expression level in an unactivated GC, where  
XX differential expression of Gs is indicative of GCA. Also included are  
XX modulating (M2) GA by contacting GC with an agent that alters the  
XX expression of at least one gene in Gs; (2) screening (M3) for an agent  
XX capable of modulating GCA or an inflammation (especially chronic) in a  
XX tissue, an allergic response in a subject, exposure of a subject to a  
XX pathogen or sterile inflammatory disease using the gene expression  
XX profile; (3) detecting (M4) an inflammation (especially chronic) in a  
XX tissue, an allergic response in a subject, exposure of a subject to a  
XX pathogen or sterile inflammatory disease, by detecting the level of  
XX expression in a sample of the tissue of gene(s) from Gs, where the level  
XX of expression of the gene is indicative of inflammation; (4) treating  
XX (M5) an inflammation (especially chronic) or in a tissue, an allergic  
XX response in a subject, exposure of a subject to a pathogen or sterile  
XX inflammatory disease, by contacting a tissue having inflammation with an  
XX agent that modulates the expression of gene(s) from Gs in the tissue. M1  
XX is useful for detecting GCA; M2 is useful for modulating GCA; M3 is useful  
XX for screening an agent capable of modulating GCA preferably in an  
XX inflammation in a tissue; M4 is useful for detecting an inflammation  
XX (especially chronic) in a tissue, an allergic response in a subject,  
XX exposure of a subject to a pathogen or sterile inflammatory disease (e.g.  
XX psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,  
XX cardiac reperfusion injury, renal reperfusion injury, ARDS, adult  
XX respiratory distress syndrome, inflammatory bowel disease, Crohn's  
XX disease, ulcerative colitis, periodontal disease, also bacterial  
XX infection, viral infection, parasitic infection, protozoal infection,  
XX fungal infection and M5 is useful for treating one of the above  
XX conditions. The present sequence represents a gene differentially  
XX expressed in granulocytes. Note: The sequence data for this patent did  
XX not form part of the printed specification, but was obtained in  
XX electronic format directly from WIPD at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
SQ Sequence 5204 BP; 1770 A; 867 C; 836 G; 1731 T; 0 U; 0 Other;

Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;  
QY 1516 TGAGCTTACGACCAATCATATGTAATGCTTCTGATTAAGCTGTCATTAAT 1575  
DB 971 TGAACCTTAACTTCCACATCACAGATATGAAGTTGGTTCAAGATTAATTAAT 1030  
QY 1576 TCTC-----TTTGAAAGCTCTGCTACTTACCAAGAGTCGCTACAGACGAAGA 1627  
DB 1031 TCTGCCCAAGACAGACCTGAATCTCTAGCTGCTAGAGGCTGACT-CAACTGAATCA 1089  
QY 1628 TGGCATTTAACGCTACTTGGAAAAATAGAAAAATGAGAACTATGAAAAATTCATGAG 1687  
DB 1090 TGGGTTTGACACGACTTGAAGTAGACCGGAGTGAATACTATGACAGTTTCATGANA 1149  
QY 1688 CAATGGTAAGCCTTACTTTTGGAAATGCT 1718  
DB 1150 AAATGGGTAAAGACTTATTTCTTGTGGCT 1180  
RESULT 6  
ADJ74588  
ID ADJ74588 standard; DNA; 5204 BP.  
XX  
AC ADJ74588;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Human intestinal fatty acid binding protein gene SEQ ID NO:12.  
XX  
KM db; gene; human; intestinal fatty acid binding protein; restenosis;  
KM coronary angioplasty; balloon coronary angioplasty;  
KM stent coronary angioplasty.  
XX  
OS Homo sapiens.  
XX  
PN WO2004015104-A1.  
XX  
PD 19-FEB-2004.  
XX  
PF 20-MAR-2003; 2003MO-JP003478.  
XX  
PR 09-AUG-2002; 2002JP-00233041.  
XX  
PA (NAGO-) NAGOYA IND SCI RES INST.  
PA (GIFU-) GIFU INT INST BIOTECHNOLOGY.  
XX  
PI Yamada Y, Yokota M;  
XX  
DR WPI; 2004-180672/17.  
XX  
PT Analysis of specific gene polymorphisms in clinical nucleic acid sample  
PT for prediction of risk of restenosis after balloon or stent coronary  
PT angioplasty.  
XX  
PS Disclosure; SEQ ID NO 12; 164pp; Japanese.  
XX  
XX The invention relates to a novel method for predicting the risk of  
XX restenosis after coronary angioplasty comprising analyzing specific gene  
XX polymorphisms in a clinical nucleic acid sample. The method is useful for  
XX the diagnosis of the genetic risk of restenosis following balloon or  
XX stent coronary angioplasty. The method has high accuracy and high  
XX estimation ratio. The present sequence is used in the exemplification of  
XX the invention.  
SQ Sequence 5204 BP; 1770 A; 867 C; 836 G; 1731 T; 0 U; 0 Other;  
Query Match 2.1%; Score 49.8; DB 12; Length 5204;  
Best Local Similarity 59.2%; Pred. No. 0.31;  
Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;  
QY 1516 TGAGCTTACGACCAATCATATGTAATGCTTCTGATTAAGCTGTCATTAAT 1575  
DB 971 TGAACCTTAACTTCCACATCACAGATATGAAGTTGGTTCAAGATTAATTAAT 1030

Qy	1576	TCTC-----TTTGAAAGCTCGTACTTACGAAAGTGTGCTTCAGACGAAAGA	1627
Db	1031	TCTGCGCCAGACAGACACCTGTAATCTTACTGCTCTAGAGGCTACT-CAACTGAATCA	1089
Qy	1628	TGGCATTTTACGGTACTTGTGAAATAATGAGAAAAATGAGACTATGAAAAATTCATGAG	1687
Db	1090	TGGGCTTTGACAGCACTTGGAAGGTGACCGGAGTGAATACTATGACAGATTCAATGAAA	1149
Qy	1688	CAATGGGTACGCTTACTTTTGTGAATGCT	1718
Db	1150	AAATGGGTAAAGACTTATTAATCTTTGTGCT	1180
RESULT 7			
ID	ADO78215	standard; DNA; 5204 BP.	
XX	ADO78215;		
AC	ADO78215;		
XX	29-JUL-2004	(first entry)	
DT			
XX			
DE	Human fatty acid binding protein 2 Alu repeat region DNA with SNP Seq 9.		
XX			
KM	human; ds; single nucleotide polymorphism; SNP; hereditary risk;		
KM	coronary spasm; tumor necrosis factor alpha;		
KM	NADH/NADPH oxidase p22 phox gene; apolipoprotein B; apolipoprotein C-III;		
KM	stromelysin 1; interleukin 6;		
KM	endothelium type nitrogen monoxide synthetase;		
KM	fatty acid binding protein 2; cardiac sudden death;		
KM	variant angina pectoris; myocardial infarction.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	variation	replace(2445..A)	
FT		/*tag= a	
XX		/standard_name= "single nucleotide polymorphism"	
PN	JP2004113093-A.		
XX			
PD	15-APR-2004.		
XX			
PF	25-SEP-2002; 2002JP-00280031.		
XX			
PR	25-SEP-2002; 2002JP-00280031.		
XX			
PA	(NAGO-) ZH NAGOYA SANGYO KAGAKU KENKYUSHO.		
PA	(GIFU-) ZH GIFUKEN KOKUSAI BIO KENYUO.		
XX			
DR	WPI; 2004-360127/34.		
XX			
PT	Detecting genotype in nucleic acid sample useful for diagnosing risk of		
PT	coronary spasm, by analyzing polymorphisms in tumor necrosis factor alpha		
PT	gene, NADH/NADPH oxidase p22 phox gene and apolipoprotein B gene, in		
PT	nucleic acid sample.		
XX			
PS	Disclosure; SEQ ID NO 9; 59pp; Japanese.		
XX			
CC	This invention relates to a novel method for detecting a hereditary risk		
CC	for a coronary spasm. Specifically, it refers to the identification of		
CC	two or more polymorphisms and determining the genotype of the nucleic		
CC	acid sample in order to calculate the risk of the patient to a coronary		
CC	spasm. The present invention describes a risk analysis that comprises		
CC	identifying two or more polymorphisms occurring at any position taken		
CC	from -863 of the tumor necrosis factor alpha gene, 242 of the NADH/NADPH		
CC	oxidase p22 phox gene, -219 of the apolipoprotein B gene, -482 of the		
CC	apolipoprotein C-III gene, -1171 of the stromelysin 1 gene, 4070 of the		
CC	apolipoprotein B gene, -634 of the interleukin 6 gene, -786 of the		
CC	endothelium type nitrogen monoxide synthetase gene or position 2445 of		
CC	the fatty acid binding protein 2 gene. Accordingly, the method enables		
CC	prevention of cardiac sudden death resulting from variant angina		
CC	pectoris, coronary spasm and myocardial infarction. This polymnucleotide		

CC	is the human fatty acid binding protein 2 gene A11 repeat region
CC	containing a G245A SNP in the sequence given in the invention. NOTE:
CC	This sequence is referred to as SeqID 9 in the sequence listing, but
CC	SeqID 8 also refers to a primer given in the disclosure of the
CC	specification.
XX	
SQ	Sequence 5204 BP; 1770 A; 867 C; 836 G; 1731 T; 0 U; 0 Other;
	Query Match            2.1%; Score 49.8; DB 12; Length 5204;
	Best Local Similarity   59.2%; Pred. No. 0.31;
	Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;
OY	1516 TGACCTTAGCGCAGCCAGCATCATATTAATTGCTTTCCTGTAAACCCTGTCATAAT 1575
DB	971 TGAACTTTAAGCTTCACATCACATGATGAAGTTGGTCAAGATMAAAAATATATATAAT 1030
OY	1576 TCCTC-----TTTGAAAGCTGTGCTACTTACCCAGAAGTCTGCCTACAGACAGAAAGA 1627
DB	1031 TCTTGCCCAAGAGACAGACCTGAATCTTCTAGCTGCTTGAAGGCTGACT-CAACTGAAATCA 1089
OY	1628 TGGCATTTAAACGCTACTTGGAAAAATAGAGAAAAATGAGAACTATGAAAAATTCATGGAG 1687
DB	1090 TGGCGTTTGACAGACACTTGGAAAGGTAGACCGGAGTGA AAACTATGACAAGTTCATGAAA 1149
OY	1688 CAATGGGTAAAGCTTACTTTTGTGAATGCCCT 1718
DB	1150 AAATGGGTAAAGACTTATATTTCTTTGTGGCT 1180
RESULT 8	
AED14682	
ID	AED14682 standard; DNA; 426 BP.
XX	
AC	AED14682;
XX	
DT	01-DEC-2005 (first entry)
DE	Mutant rat intestinal fatty acid binding protein DNA.
XX	
KW	high throughput screening; diagnosis; fluorescence; drug metabolism;
KM	cerebrovascular ischemia; cerebroprotective; vasodilator; heart disease;
KM	cardiac; cardiovascular disease; cancer; cytostatic; neoplasm;
XX	intestinal fatty acid binding protein; I-FABP; mutant; ds; gene.
OS	Rattus norvegicus.
OS	Synthetic.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..423
FT	/*tag= a
FT	/product= "Intestinal fatty acid binding protein"
XX	
PN	WO2005093103-A2.
XX	
PD	06-OCT-2005.
XX	
PF	21-MAR-2005; 2005WO-US009278.
XX	
PR	22-MAR-2004; 2004US-0555224P.
PA	(FPA)- FPA SCI LLP.
XX	
P1	Kleinfeld AM, Huber AH, Kampf JP, Kwan T, Zhu B;
DR	WPI; 2005-713873/73.
DR	P-PSDB; AED14683.
XX	
PT	High throughput generation and screening of probes to determine
PT	concentration of unbound bilirubin in, e.g., blood, comprises generating
PT	polynucleotides, expressing proteins, purifying, associating with
PT	fluorophores, retrieving and screening.
XX	
SS	Disclosure; SEQ ID NO 3; 82pp; English.









CC from wipo.int/pub/published\_pct\_sequences  
XX Sequence 2636 BP; 849 A; 414 C; 497 G; 876 T; 0 U; 0 Other;  
SQ Sequence 2636 BP; 849 A; 414 C; 497 G; 876 T; 0 U; 0 Other;  
Query Match 2.0%; Score 47; DB 4; Length 2636;  
Best Local Similarity 52.3%; Pred. No. 1.1;  
Matches 104; Conservative 0; Mismatches 95; Indels 0; Gaps 0;  
QY 1821 ATTGCGCTGCACCTATTGTAAGAAAGCTCTATAGAGGGAAATACAGAGAAACAAATT 1880  
DB 540 ATTGACTTCGAATATCTTTAAATAATACAGCCCTGTTGCTTCTCGAATTTAAACAT 599  
QY 1881 CTGATTTATTTCGACTTCGCATCTTATGCAATTAGCTAATTCAGTAGAGCAATCC 1940  
DB 600 ATAAATTATTTTAAATGCTGAATTAATTCCTCTTAAATAATTCAGGTATATTATTTC 659  
QY 1941 AGCAGAAATTTAATAGATTATAGTATAGAAATTTATTTTGAATAGACTTTGAAA 2000  
DB 660 TTAGGGAAATTAAGTATTATTTGCTGCTAGTTTATATGTTAAAAAAGTATAGACAA 719  
QY 2001 ATTACACAAAGAGGAAATT 2019  
DB 720 ATTAGTAGATGTGTAATTT 738  
RESULT 14  
ADD29587  
ID ADD29587 standard; mRNA; 662 BP.  
XX  
AC ADD29587;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Mouse tumour suppressor mRNA SEQ ID NO:36.  
XX  
KM 88; mouse; tumour suppressor; cancer; cancer; cytostatic; gene therapy.  
XX  
OS Mus musculus.  
XX  
PN W02003058201-A2.  
XX  
PD 17-JUL-2003.  
XX  
PF 31-DEC-2002; 2002W0-US041825.  
XX  
PR 31-DEC-2001; 2001US-0345317P.  
XX  
PA (QUAR-) QUARK BIOTRCH INC.  
XX  
PA (CLEV-) CLEVELAND CLINIC FOUND.  
XX  
PI Feinstein E, Gudkov AV;  
XX  
DR WPI; 2003-598393/56.  
XX  
PT Diagnosing cancer comprises determining the polypeptide or polynucleotide  
PT levels e.g., hepatic lipase, in a sample from a subject, where a higher  
PT level compared to that in a subject free of cancer is indicative of  
PT cancer.  
XX  
PS Disclosure; SEQ ID NO 36; 272pp; English.  
XX  
XX The invention relates to a novel method for diagnosing a cancer in a  
CC subject. the method comprises determining, in a sample from the cancer in a  
CC the level of at least one polypeptide, where a higher level of the  
CC polypeptide compared to the level of the polypeptide in a subject free of  
CC cancer is indicative of cancer. The polypeptide is selected from any of  
CC the polypeptides encoded by the polynucleotides listed in the  
CC specification and polypeptides which are at least 70% homologous to the  
CC polypeptides. The method of the invention has cytostatic activity, and  
CC may have a use in gene therapy. The method is useful in identifying  
CC markers specific for one or several types of cancer, depending on the  
CC tissue origin, which may be used in numerous diagnostic and prognostic  
CC applications as well as cancer type-specific targets for therapeutic

CC intervention. The compounds that modulate the activity of a tumour  
CC suppressor gene are useful in the treatment of cancer or as anti-cancer  
CC drugs. The present sequence represents a polynucleotide of the invention.  
XX  
SQ Sequence 662 BP; 217 A; 133 C; 148 G; 164 T; 0 U; 0 Other;  
Query Match 2.0%; Score 46.8; DB 10; Length 662;  
Best Local Similarity 73.2%; Pred. No. 0.8;  
Matches 60; Conservative 0; Mismatches 22; Indels 0; Gaps 0;  
QY 1615 ACAGACAAAAGATGGCATTTAAGCTGCTTGGAAAAATAGAGAAAATGAGAACTATGAA 1674  
DB 28 ACAGCTGATCATGCGATTCGACCGCACGTGAAAGTAGACCCGAAACGAAACTATGAA 87  
QY 1675 AAATTCATGAGCAATGCGTA 1696  
DB 88 AAGTTCATGAGAAAATGGCA 109  
RESULT 15  
ADZ62521  
ID ADZ62521 standard; DNA; 670 BP.  
XX  
AC ADZ62521;  
XX  
DT 30-JUN-2005 (first entry)  
XX  
DE Murine Fabp2 gene, SEQ ID 2245.  
XX  
KM Drug screening; Peroxisome Proliferator-Activated Receptor alpha;  
KM PPAR alpha; gene; db.  
XX  
OS Mus musculus.  
XX  
PN US2005084872-A1.  
XX  
PD 21-APR-2005.  
XX  
PF 23-JAN-2004; 2004US-00764420.  
XX  
PR 24-JAN-2003; 2003US-0442797P.  
XX  
PR 30-MAY-2003; 2003US-0474413P.  
XX  
PA (LUMP/) LUM P Y.  
PA (TANY/) TAN Y.  
PA (DAIH/) DAI H.  
PA (MUIS/) MUISE E S.  
PA (BERG/) BERGER J P.  
PA (THOM/) THOMPSON J R.  
XX  
PI Lum PY, Tan Y, Dai H, Muise ES, Berger JP, Thompson JR;  
XX  
DR WPI; 2005-313963/32.  
XX  
DR REFSQ; NM\_007980.  
XX  
PT Determining whether agent has defined biological activity, by making  
PT comparison e.g. comparing efficacy value, toxicity value and classifier  
PT value of agent to reference values, and using results to determine  
PT whether agent has activity.  
XX  
PS Example 4; SEQ ID NO 2245; 51pp; English.  
XX  
XX The present invention relates to a method for screening biologically  
CC active agents, such as candidate drugs, to determine whether the agent  
CC possesses a defined biological activity. The methods involve making one  
CC or more comparisons chosen from comparing efficacy value of agent to  
CC reference efficacy value, comparing a toxicity value of the agent to  
CC reference toxicity value, comparing a classifier value of the agent to  
CC reference classifier value, and using the comparison result(s) to  
CC determine whether the agent possesses the defined biological activity.  
CC Also claimed is a population of oligonucleotide probes (I) specific for  
CC measuring the expression levels of members of a classifier population of  
CC genes or a toxicity-related population of genes. (I) are useful for

CC measuring the expression levels of genes that are useful for identifying  
 CC agonists or partial agonists of Peroxisome Proliferator-Activated  
 CC Receptor (PPAR) gamma. ADZ63072-ADZ63959 are oligonucleotide probes which  
 CC are useful in the method of the invention to measure the expression  
 CC pattern of mouse genes ADZ62273-ADZ63071 that yield an expression pattern  
 CC that correlates with the stimulation of PPARalpha receptors by an agent.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from USPTO  
 CC at seqdata.uspto.gov/sequence.html?DocID=20050084872.

CC  
 XX  
 SQ Sequence 670 BP; 228 A; 131 C; 147 G; 164 T; 0 U; 0 Other;

Query Match 2.0%; Score 46.8; DB 14; Length 670;  
 Best Local Similarity 73.2%; Pred. No. 0.81;  
 Matches 60; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1615 ACAGACAGAAAGATGGCATTTAACGCTACTTGGAAAAATGAGAACTATGAA 1674  
 Db 24 ACAGCTGAGATCATGCGATTGACGCGACGTGGAAAGTAAACCGAACGAACTATGAA 83

QY 1675 AAATTGATGGAAGCATGGGTA 1696  
 Db 84 AAGTTGATGGAAGAAATGGGCA 105

Search completed: May 27, 2006, 19:15:54  
 Job time : 1601.66 secs

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: May 28, 2006, 05:53:57 ; Search time 12207.9 Seconds  
(without alignments)  
10906.400 Million cell updates/sec

Title: US-10-099-663-1

Perfect score: 2381  
Sequence: 1 agctccctgcgcagaaag.....atctgaagctcaccatctcag 2381

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est3:\*  
3: gb\_est4:\*  
4: gb\_est5:\*  
5: gb\_est6:\*  
6: gb\_est7:\*  
7: gb\_est8:\*  
8: gb\_est9:\*  
9: gb\_est10:\*  
10: gb\_est11:\*  
11: gb\_est12:\*  
12: gb\_est13:\*  
13: gb\_est14:\*  
14: gb\_est15:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	752.6	31.6	1473	12	CC250935 CH261-17C
2	99	4.2	598	5	CF250818 esa017_e0
3	95	4.0	595	5	CF250947 esa019_a0
4	85	3.6	705	3	BU255106 60347428
5	84	3.5	759	3	BU265304 603508640
6	81	3.4	829	3	BU296697 603741965
7	79.4	3.3	885	3	BU123280 603003274
8	79	3.3	814	3	BU234099 603792609
9	78.8	3.3	1080	3	BU123336 603149475
10	75.8	3.2	426	5	CD739695 4028031_1
11	74	3.1	576	5	CD739141 4028481_1
12	70.4	3.0	480	5	CD733346 4045132_1
13	62.4	2.6	1101	14	CNS00396 AL063921 Drosophila
14	61.4	2.6	885	3	BU123717 603147391
15	60.4	2.5	443	4	BK276255 BX276255
16	57.2	2.4	683	8	CW793516 4128553_B
17	55.8	2.3	657	5	CK945920 4070423_B
18	55	2.3	730	10	DT824420 LB00233.C
19	55	2.3	731	10	DT824909 LB00235.C

20	54.2	2.3	572	5	CK834687 4059069_B
21	54.2	2.3	673	5	CK947902 4072726_B
22	54.2	2.3	686	5	CK957350 4097834_B
23	54.2	2.3	708	5	CK960577 4101917_B
24	54	2.3	481	2	BM430667 1DU003A11
25	54	2.3	489	4	CB224035 1JB370F5
26	54	2.3	497	2	BM430956 1DU007F03
27	54	2.3	548	2	BM432416 1JEC4B4_A
28	54	2.3	568	2	BM431104 1DU009C09
29	54	2.3	671	8	CK980160 4111791_B
30	54	2.3	679	5	CK947198 4071842_B
31	54	2.3	682	5	CK948900 4074042_B
32	54	2.3	684	5	CK947048 4071436_B
33	54	2.3	702	5	CK944155 4068182_B
34	54	2.3	720	5	CK946512 4071131_B
35	53.4	2.2	669	10	DT820486 LB00214.C
36	53.4	2.2	726	10	DT820499 LB00021.C
37	53.4	2.2	726	10	DT821139 LB00021.C
38	53.4	2.2	726	10	DT822486 LB00024.C
39	53.4	2.2	726	10	DT825113 LB000236.C
40	53.4	2.2	727	10	DT819495 LB000210.C
41	53.4	2.2	727	10	DT823340 LB000228.C
42	53.4	2.2	823	10	DT822459 LB00022.C
43	53.4	2.2	843	10	DT828569 LB00247.C
44	53	2.2	372	2	BM430450 1DU034E2
45	53	2.2	414	2	BM433066 1JEC9E2.A

#### ALIGNMENTS

RESULT 1  
LOCUS CC250935/1473 bp DNA linear GSS 13-MAY-2003  
DEFINITION CH261-17C7\_Sp6.1 CH261 Gallus gallus genomic clone CH261-17C7,  
genomic survey sequence.

ACCESSION CC250935  
VERSION CC250935.1 GI:30587685  
KEYWORDS GSS.  
SOURCE Gallus gallus (chicken)  
ORGANISM Gallus gallus

REFERENCE 1 (bases 1 to 1473)  
AUTHORS Kremetzki,C., Higgindobham,J., Wylie,K., Carter,J., McPherson,J.,  
Archoseauria,Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.

TITLE Warren,W., Graves,T., Wardis,E. and Wilson,R.  
JOURNAL Gallus gallus BAC End Reads  
COMMENT Unpublished (2003)  
Contact: Richard K. Wilson  
Genome Sequencing Center  
Washington University School of Medicine  
Email: submissions@wustl.edu  
Insert Length: 182000 Std Error: 0.00  
Seq primer: Sp6 ATTACGTGACACTATAG  
Class: BAC ends  
High quality sequence start: 239  
High quality sequence stop: 912.  
Location/Qualifiers

#### FEATURES

source

1. 1473  
/organism="Gallus gallus"  
/mol\_type="genomic DNA"  
/strain="Red Jungle Fowl"  
/db\_xref="taxon:9031"  
/clone="CH261-17C7"  
/sex="female"  
/cell\_line="UCD001, inbred 256"  
/clone\_1bp="CH261"  
/note="Vector: pPARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;  
CH261 Female Chicken library - For library and clone  
ordering information: http://www.chori.org/bacpac"

ORIGIN

Query Match	Best Local	Similarity	31.6%	Score	752.6	DB 12	Length	1473	
Matches	793	Conservative	0	Mismatches	9	Indels	5	Gaps	3
Qy	1580	TTTGCAAGCTCTGCTACTTACCCAGAAAGTC-TGCTTACAGACAGAAAGATGGCATTTAA-	1637						
Db	1042	TTGCAAAAGTTGCTTCTTACAGAAATCTTCCCTACAGACAGAAAGATGGCATTTAAAC	983						
Qy	1638	CGGTACTTGGAAATATGAGAAAAATGAGAACTATGAAAAATTCATGAAAGCAATGGCTTA	1697						
Db	982	CGGTACTTGGAAATATGAGAAAAATGAGAACTATGAAAAATTCATGAAAGCAATGGCTTA	923						
Qy	1698	GCCTTACTTTTTTGAATGCGCTTCTAAAGCAGGATACCACTACGGCGGAATCAACTTA	1757						
Db	922	GCCTTACTTTTTTGAATGCGCTTCTAAAGCAGGATACCACTACGGCGGAATCAACTTA	863						
Qy	1758	AGCTGTTCAATGAACTACCTCTGGCTAACCTGTCTCTTTGTTGTCTATATTTGCGCTTG	1817						
Db	862	AGCTGTTCAATGAACTACCTCTGGCTAACCTGTCTCTTTGTTGTCTATATTTGCGCTTG	803						
Qy	1818	CACATTCGCCCTGCCTACTTATTTTGAAGAAAGCTCATATGAGGGGAAATCAAGAAAGAAAC	1877						
Db	802	CACATTCGCCCTGCCTACTTATTTTGAAGAAAGCTCATATGAGGGGAAATCAAGAAAGAAAC	743						
Qy	1878	ATTTCGATTTTATTTGCAATTCGCATATCTTATGCAATTTAGCTAATTCAGTAGAGCAT	1937						
Db	742	ATTTCGATTTTATTTGCAATTCGCATATCTTATGCAATTTAGCTAATTCAGTAGAGCAT	683						
Qy	1938	TCCAGACAGAAATTTAAATATGAAATATATGTAAGAAATATTTTGTATATGAACTGTTGA	1997						
Db	682	TCCAGACAGAAATTTAAATATGAAATATATGTAAGAAATATTTTGTATATGAACTGTTGA	623						
Qy	1998	AAAAATTACACAAGAGGAAATTCGTGCTCCAGTTTTCAGAAACAACATGATTTGAGT	2057						
Db	622	AAAAATTACACAAGAGGAAATTCGTGCTCCAGTTTTCAGAAACAACATGATTTGAGT	563						
Qy	2058	CATTTTAACATGCTAGTGTCTTACTTTAAGCTTTGTAACAACTGCTGTATATGATGTAA	2117						
Db	562	CATTTTAACATGCTAGTGTCTTACTTTAAGCTTTGTAACAACTGCTGTATATGATGTAA	503						
Qy	2118	ACATACTATCTCTAGTGTGATAGTATGTTGTAATPACAGGCTGAAACACTGCTCAGTGA	2177						
Db	502	ACATACTATCTCTAGTGTGATAGTATGTTGTAATPACAGGCTGAAACACTGCTCAGTGA	443						
Qy	2178	GGTGAGAAAGTAAAGACTCTGAGTCAGAAATTCGTGGCTAAGCTCCTCAACTACAGAA	2237						
Db	442	GGTGAGAAAGTAAAGACTCTGAGTCAGAAATTCGTGGCTAAGCTCCTCAACTACAGAA	383						
Qy	2238	AAGTCACATATAAAAATGCAACATGATGTCATATTTGTTTCTCTGCTGATGTTAAT	2297						
Db	382	AAGTCACATATAAAAATGCAACATGATGTCATATTTGTTTCTCTGCTGATGTTAAT	323						
Qy	2298	TGATTTATTATTA--TTTTTTTTTTAGGCGTGAATGTATGAAAGAAATTAGAGCC	2354						
Db	322	TGATTTATTATTAATTTTTTTTTTTTTTTTGTAGCGCTGAATGTATGAAAGAAATTAGAGCC	263						
Qy	2355	CACGATATCTGAAGCTCAGTATTCAG	2381						
Db	262	CACGATATCTGAAGCTCAGTATTCAG	236						
RESULT 2	CF250818	598 bp	mRNA	linear	EST 07-AUG-2003				
LOCUS	CF250818								
DEFINITION	caeo01_095 Bimejiria tenella-infected caecal tonsil Gallus gallus								
ACCESSION	CF250818								
VERSION	CF250818.1								
KEYWORDS	EST								
SOURCE	Gallus gallus (chicken)								
ORGANISM	Gallus gallus								
	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae;								

REFERENCE Phaseianinae; Gallus.  
AUTHORS 1 (bases 1 to 598)  
TITLE Wiltzell,H., Bed'Hom,B., Morin,V., Young,J.R., Whitaker,C.J.,  
JOURNAL Chausse,A.M. and Zoorob,R.  
COMMENT A collection of chicken ESTs from activated immune cells  
Unpublished (2003)  
Contact: Zoorob R  
UPR 1983  
CNRS

FEATURES  
source 7 rue Guy Moquet, BP 8, 94801 Villejuif cedex, France  
Tel: 33 1 49 58 35 00  
Fax: 33 1 49 58 33 81  
Email: zoorob@vjf.cnrs.fr.  
Location/Qualifiers  
1. 598  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9031"  
/clone\_lib="Bimelirix tenella-infected caecal tonsil"  
/note="Organ: Caecal tonsil; Vector: pT1pLEX2"

ORIGIN  
Query Match 4.2%; Score 99; DB 5; Length 598;  
Best Local Similarity 100.0%; Pred. No. 2,7e-12;  
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1596 ACTTACCGAAGTCTGCTACAGACAGAAAGATGSCATTTAACGCTACTTGGAAAAATGA 1655  
Db 3 ACTTACCGAAGTCTGCTACAGACAGAAAGATGSCATTTAACGCTACTTGGAAAAATGA 62

QY 1656 GAAAAATGAGAACTATGAAAAAATTCATGAGCAATGGG 1694  
Db 63 GAAAAATGAGAACTATGAAAAAATTCATGAGCAATGGG 101

RESULT 3  
LOCUS CF250947 595 bp mRNA linear EST 07-AUG-2003  
DEFINITION esa019 a07 Bimelirix tenella-infected caecal tonsil Gallus gallus  
cdna, mRNA sequence.  
ACCESSION CF250947  
VERSION CF250947.1 GI:33484202  
KEYWORDS EST.  
SOURCE Gallus gallus (chicken)  
ORGANISM Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archaeoptera; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
1 (bases 1 to 595).  
Wiltzell,H., Bed'Hom,B., Morin,V., Young,J.R., Whitaker,C.J.,  
Chausse,A.M. and Zoorob,R.  
A collection of chicken ESTs from activated immune cells  
Unpublished (2003)  
Contact: Zoorob R  
UPR 1983  
CNRS

REFERENCE 7 rue Guy Moquet, BP 8, 94801 Villejuif cedex, France  
AUTHORS Tel: 33 1 49 58 35 00  
TITLE Fax: 33 1 49 58 33 81  
JOURNAL Email: zoorob@vjf.cnrs.fr.  
COMMENT Unpublished (2003)  
Contact: Zoorob R  
UPR 1983  
CNRS

FEATURES  
source 7 rue Guy Moquet, BP 8, 94801 Villejuif cedex, France  
Tel: 33 1 49 58 35 00  
Fax: 33 1 49 58 33 81  
Email: zoorob@vjf.cnrs.fr.  
Location/Qualifiers  
1. 595  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9031"  
/clone\_lib="Bimelirix tenella-infected caecal tonsil"  
/note="Organ: Caecal tonsil; Vector: pT1pLEX2"

ORIGIN  
Query Match 4.0%; Score 95; DB 5; Length 595;  
Best Local Similarity 100.0%; Pred. No. 2.5e-11;  
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1600 ACCAGAGCTCTGCTACAGACAGAAAGATGCGATTAAACGTAATTGGAAAAATAGAGAAA 1659  
 Db 3 ACCAGAGCTCTGCTACAGACAGAAAGATGCGATTAAACGTAATTGGAAAAATAGAGAAA 62  
 QY 1660 AATGGAACCTATGAAAAATTCATGAGCAATGGG 1694  
 Db 63 AATGGAACCTATGAAAAATTCATGAGCAATGGG 97

RESULT 4  
 BU355106  
 LOCUS 705 bp mRNA linear EST 28-NOV-2002  
 DEFINITION 603474288P1 CSEQCHN70 Gallus gallus CDNA clone CHEST355022 5', mRNA  
 sequence.

ACCESSION BU355106  
 VERSION BU355106.1 GI:25863107  
 KEYWORDS EST.  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 705)  
 AUTHORS Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,  
 Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.  
 TITLE A Comprehensive Collection of Chicken CDNAs  
 JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)  
 PUBMED 12445392

COMMENT Contact: Simon Hubbard  
 Department of Biomolecular Sciences  
 University of Manchester Institute of Science and Technology  
 (UMIST)  
 PO Box 88, Manchester, M60 1QD, UK  
 Tel: 01612008930  
 Fax: 01612360409  
 Email: Simon.Hubbard@umist.ac.uk.

FEATURES  
 source location/Qualifiers

1..705  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /strain="White Leghorn, Hisex"  
 /db\_xref="taxon:9031"  
 /clone="CHEST355022"  
 /dev\_stage="36"  
 /lab\_host="DH10B"  
 /clone\_lib="CSEQCHN70"  
 /note="Organ: hearts; Vector: pBluescript II KS(+); Site\_1: EcoRI; Site\_2: NotI; This normalized library was constructed from 1 million independent clones. CDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was bluntended, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN

Query Match 3.6%; Score 85; DB 3; Length 705;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-09;  
 Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1610 TGCCCTACAGACAGAAAGATGCGATTAAACGTAATTGGAAAAATAGAGAAAATGAGAACT 1669  
 Db 1 TGCCCTACAGACAGAAAGATGCGATTAAACGTAATTGGAAAAATAGAGAAAATGAGAACT 60  
 QY 1670 ATGAAAAATTCATGAGCAATGGG 1694  
 Db 61 ATGAAAAATTCATGAGCAATGGG 85

RESULT 5  
 BU265304  
 LOCUS 759 bp mRNA linear EST 26-NOV-2002  
 DEFINITION 603508640P1 CSEQCHN52 Gallus gallus CDNA clone CHEST437010 5', mRNA  
 sequence.

ACCESSION BU265304  
 VERSION BU265304.1 GI:25536254  
 KEYWORDS EST.  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 759)  
 AUTHORS Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,  
 Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.  
 TITLE A Comprehensive Collection of Chicken CDNAs  
 JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)  
 PUBMED 12445392

COMMENT Contact: Simon Hubbard  
 Department of Biomolecular Sciences  
 University of Manchester Institute of Science and Technology  
 (UMIST)  
 PO Box 88, Manchester, M60 1QD, UK  
 Tel: 01612008930  
 Fax: 01612360409  
 Email: Simon.Hubbard@umist.ac.uk.

FEATURES  
 source location/Qualifiers

1..759  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /strain="White Leghorn, Hisex"  
 /db\_xref="taxon:9031"  
 /clone="CHEST437010"  
 /dev\_stage="22"  
 /lab\_host="DH10B"  
 /clone\_lib="CSEQCHN52"  
 /note="Organ: limbs; Vector: pBluescript II KS(+); Site\_1: EcoRI; Site\_2: NotI; This normalized library was constructed from 1 million independent clones. CDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was bluntended, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN

Query Match 3.5%; Score 84; DB 3; Length 759;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-08;  
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1611 GCCTACAGACAGAAAGATGCGATTAAACGTAATTGGAAAAATAGAGAAAATGAGAACTA 1670  
 Db 1 GCCTACAGACAGAAAGATGCGATTAAACGTAATTGGAAAAATAGAGAAAATGAGAACTA 60  
 QY 1671 TGAATAATTCATGAGCAATGGG 1694  
 Db 61 TGAATAATTCATGAGCAATGGG 84

RESULT 6  
 BU296697  
 LOCUS 829 bp mRNA linear EST 27-NOV-2002  
 DEFINITION 603741965P1 CSEQCHN56 Gallus gallus CDNA clone CHEST640b17 5', mRNA  
 sequence.

ACCESSION BU296697  
 VERSION BU296697.1 GI:25746333  
 KEYWORDS EST.  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.  
 1 (bases 1 to 829)  
 Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Pong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.  
 A Comprehensive Collection of Chicken cDNAs  
 Curr. Biol. 12 (22), 1965-1969 (2002)  
 12445392  
 CONTACT: Simon Hubbard  
 Department of Biomolecular Sciences  
 University of Manchester Institute of Science and Technology (UMIST)  
 PO Box 88, Manchester, M60 1QD, UK  
 Tel: 01612008930  
 Fax: 01612360409  
 Email: Simon.Hubbard@umist.ac.uk.  
 Location/Qualifiers  
 1..829  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /strain="Compton Line 151"  
 /db\_xref="taxon:9031"  
 /clone="CHEST640b17"  
 /sex="Female"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_1lb="CSEQCHN56"  
 /note="Organ: small intestine; Vector: pBluescript II KS(+); Site\_1: EcoRI; Site\_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunt-ended, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1996): 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN

Query Match 3.4%; Score 81; DB 3; Length 829;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-08;  
 Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1614 TACAGACAGAAAGATGGCATTTAAACGTAAGTGAAGAAATGAGAAATGAACTATGA 1673  
 |||||||  
 DB 11 TACAGACAGAAAGATGGCATTTAAACGTAAGTGAAGAAATGAGAAATGAACTATGA 70  
 |||||||  
 QY 1674 AAAATTCATGGAAGCAATGGG 1694  
 |||||||  
 DB 71 AAAATTCATGGAAGCAATGGG 91  
 |||||||

RESULT 7  
 BU123280 885 bp mRNA linear EST 25-NOV-2002  
 LOCUS 603003274F1 CSEQCHL18 Gallus gallus cDNA clone CHEST14C20 5', mRNA  
 DEFINITION sequence.  
 ACCESSION BU123280  
 VERSION BU123280.1 GI:25333903  
 KEYWORDS EST.  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.  
 1 (bases 1 to 885)  
 Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Pong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.  
 A Comprehensive Collection of Chicken cDNAs  
 Curr. Biol. 12 (22), 1965-1969 (2002)  
 12445392  
 CONTACT: Simon Hubbard  
 Department of Biomolecular Sciences  
 University of Manchester Institute of Science and Technology (UMIST)  
 PO Box 88, Manchester, M60 1QD, UK  
 Tel: 01612008930  
 Fax: 01612360409  
 Email: Simon.Hubbard@umist.ac.uk.  
 Location/Qualifiers  
 1..885  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /strain="Compton Line 151"  
 /db\_xref="taxon:9031"  
 /clone="CHEST14C20"  
 /sex="Female"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_1lb="CSEQCHL18"  
 /note="Organ: small intestine; Vector: pBluescript II KS(+); Site\_1: EcoRI; Site\_2: NotI; Modification of pBluescript II KS(+) (Stratagene) vector to accommodate cDNA produced with the T-timed protocol (Construction of uni-directionally cloned cDNA libraries from messenger RNA for improved 3' end DNA sequencing by Glenn Fu, et al. U.S. Patent # 6,387,624). Cut pBluescript II KS(+) with NotI and EcoRI. Ligate in double stranded adaptor containing BspGI and BamHI sites [5'ggccgctgcagcccgatccggaataaag] [5'aattcttttttcggtccgcgcgcgc]"

ORIGIN

Query Match 3.3%; Score 79.4; DB 3; Length 885;  
 Best Local Similarity 98.8%; Pred. No. 1.4e-07;  
 Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1614 TACAGACAGAAAGATGGCATTTAAACGTAAGTGAAGAAATGAGAAATGAACTATGA 1673  
 |||||||  
 DB 1 TACAGACAGAAAGATGGCATTTAAACGTAAGTGAAGAAATGAGAAATGAACTATGA 60  
 |||||||  
 QY 1674 AAAATTCATGGAAGCAATGGG 1694  
 |||||||  
 DB 61 AAAATTCATGGAAGCAATGGG 81  
 |||||||

RESULT 8  
 BU234099 814 bp mRNA linear EST 26-NOV-2002  
 LOCUS 603792609F1 CSEQCHN24 Gallus gallus cDNA clone CHEST758B4 5', mRNA  
 DEFINITION sequence.  
 ACCESSION BU234099  
 VERSION BU234099.1 GI:25478348  
 KEYWORDS EST.  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.  
 1 (bases 1 to 814)  
 Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Pong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.  
 A Comprehensive Collection of Chicken cDNAs  
 Curr. Biol. 12 (22), 1965-1969 (2002)  
 12445392  
 CONTACT: Simon Hubbard



Department of Biomolecular Sciences  
University of Manchester Institute of Science and Technology  
(UMIST)  
PO Box 88, Manchester, M60 1QD, UK  
Tel: 016127508930  
Fax: 01612360409  
Email: Simon.Hubbard@umist.ac.uk

**FEATURES**

**SOURCE**

```

1. .814
   /organism="Gallus gallus"
   /mol_type="mRNA"
   /strain="White Leghorn, H1sex"
   /db_xref="taxon:9031"
   /clone="C85T758b4"
   /dev_stage="22"
   /lab_host="DH10B"
   /clone_1lb="C85CGMN24"
   /note="Organ: heads; Vector: pBluescript II KS(+); Site_1
   /Site_2: NotI; This normalized library was
   constructed from 1 million independent clones; cDNA
   synthesis was initiated using an oligo(dT) primer, using
   methylated C in the first strand synthesis reaction.
   Following this first strand reaction, double-stranded cDNA
   was blunted, ligated to NotI adapters, digested with
   EcoRI, size-selected, and cloned into the NotI and EcoRI
   compatible sites of a custom modified MCS of the
   pBluescript (KS+) vector. The library was normalized in 2
   rounds using conditions adapted from Soares et al., PNAS
   (1994) 91: 9228-9232 and Bonaldo et al., Genome Research
   (1996): 791, except that a significantly longer
   reannealing hybridization was used."

```

## ORIGIN

Query Match	3.3%	Score 79;	DB 3;	Length 814;
Best Local Similarity	100.0%	Pred. No. 1.7e-07;		
Matches 79;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

Qy	1616	CAGCAGAAATGCGCATTTTAACGGTACTTGGAATAATGAGAAAATGAGAACTATGAAA	1675
Db	1	CAGCAGAAAGATGCGCATTTTAACGGTACTTGGAATAATGAGAAAATGAGAACTATGAAA	60
Qy	1676	AATTCATGGAAGCAATGGG	1694
Db	61	AATTCATGGAAGCAATGGG	79

RESULT 9	LOCUS	DEFINITION
BUI23336	BUI23336	1080 bp mRNA linear EST 25-NOV-2002
603144475F1	CSEBCHL18	Gallus gallus CDNA clone CHSST153fl19 5', mRNA sequence.

REFERENCE	1 (bases 1 to 1080)
AUTHORS	Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Butt, D.W., Bosch, E., Pong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
TITLE	A Comprehensive Collection of Chicken cDNAs
JOURNAL	Curr. Biol. 12 (22), 1965-1969 (2002)
PUBMED	12445392
COMMENT	Contact: Simon Hubbard

**FEATURES**  
**SOURCE**

```

Location/Qualifiers
1..1080
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton Line 151"
/db_xref="taxon:9031"
/clone="CHEST153f19"
/sex="Female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1fb="CSQCHL18"
/note="Organ: small intestine; Vector: pBluescript II
KS(+); Site 1: EcoRI; Site 2: NotI; Modification of
pBluescript II KS1(+). (Stratagene) vector to accommodate
cDNA produced with the T-primed protocol (Construction of
uni-directionally cloned cDNA libraries from messenger RNA
for improved 3' end DNA sequencing by Glenn Fu, et al.
U.S. Patent # 6,387,624). Cut pBluescript II KS(+) with
NotI and EcoRI. Ligate in double stranded adaptor
containing BspGI and BamHI sites
[5'ggcgcgtgcagcccccgcagtcgcgaataaag]
[5'aattctttttccggatccggggccgcgcacgc]"

```

## ORIGIN

Query Match	3.3%	Score 78.8;	DB 3;	Length 1080;
Best Local Similarity	97.6%	Pred. No. 2e-07;		
Matches 80; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0

Qy	1613	CTAAGACAGAAAGATGGCATTTTAAACGTACTTGAGAAATATGAGAAAAATGGAACATATG	1672
Db	7	CGAGAGACAGAAAGATGGCATTTTAAACGTACTTGAGAAATATGAGAAAAATGGAACATATG	66
Qy	1673	AAAAATTCATGAGAGCAATGGG	1694
Db	67	AAAAATTCATGAGAGCAATGGG	88

RESULT 10					
CD739695					
LOCUS	CD739695	426 bp	mRNA	linear	EST 26-JUN-2003
DEFINITION	4288031 IGAL - Chicken Intestinal Lymphocyte Gallus gallus CDNA				
	clone IGAL_20B02 5', mRNA sequence.				

REFERENCE	1 (Pages 1 to 426)
AUTHORS	Min, W., Lillehoj, H. S., Ashwell, C. M., Matukumalli, L. K., van Tassel, C. and Han, J. Y.
TITLE	Chicken intestinal lymphocyte ESR database as a resource for the analysis of mucosal immune function
JOURNAL	Unpublished (2003)
COMMENT	Contact: Hyun S. Lillehoj

## FEATURES

```
/organism="Gallus gallus"  
/mol_type="mRNA"
```

/strain="white leghorn SC"  
/db\_xref="taxon:9831"  
/clone="IGAL\_20802"  
/sex="mixed"  
/tissue\_type="Gut"  
/cell\_type="Lymphocyte"  
/dev\_stage="Adult"  
/lab\_host="EMDH10B"  
/clone\_lib="IGAL - Chicken Intestinal Lymphocyte"  
/note="Organ: Intestine; Vector: PCMV-SPORT6; Site 1:  
Sal; Site 2: NoCI; Normalized library from chicken gut  
infected with coccidia duodenum and middle gut."

Query Match 3.2%; Score 75.8; DB 5; Length 426;  
Best Local Similarity 97.5%; Pred. No. 8.5e-07;  
Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1616 CAGCAGAAAGATGGCTTTAAGCTTACCTTGAGAAATAGAGAAATGAGAACTATGAA 1675  
|||||  
DB 2 CAGCAGAAAGATGGCTTTAAGCTTACCTTGAGAAATAGAGAAATGAGAACTATGAA 61  
QY 1676 AATTCATGAGCAATGGG 1694  
|||||  
DB 62 AATTCATGAGCAATGGG 80

RESULT 11  
CD739141 576 bp mRNA linear EST 26-JUN-2003  
LOCUS 4026481 IGAL - Chicken Intestinal Lymphocyte Gallus gallus CDNA  
DEFINITION clone IGAL\_93A12 5', mRNA sequence.  
CD739141  
CD739141.1 GI:32289990  
EST.  
SOURCE Gallus gallus (chicken)  
ORGANISM Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
1 (bases 1 to 576)  
Min.W., Lillehoj,H.S., Ashwell,C.M., Matukumalli,L.K., van  
Tassel,C. and Han,J.Y.  
Chicken Intestinal Lymphocyte EST database as a resource for the  
analysis of mucosal immune function  
Unpublished (2003)  
Contact: Hyun S. Lillehoj  
Animal Parasite Diseases Laboratory  
Animal and Natural Resources Institute, USDA  
Bldg.1043, BARC-East, Beltsville, MD 20705, USA  
Tel: 3015048771  
Fax: 3015045103

REFERENCE Email: lillehoj@nri.barc.usda.gov  
AUTHORS Single pass sequencing. Bases called and trimmed with phred  
0.000925 using options -trim\_alt '' -trim\_faasta. Vector identified  
by cross\_match using options -mismatch 12 -minscore 12  
Plate: 93 row: A column: 12  
Seq primer: ATTTAGGTGACACTATAG  
High quality sequence stop: 576.  
Location/Qualifiers  
1..576

FEATURES  
source  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/strain="white leghorn SC"  
/db\_xref="taxon:9031"  
/clone="IGAL\_93A12"  
/sex="mixed"  
/tissue\_type="Gut"  
/cell\_type="Lymphocyte"  
/dev\_stage="Adult"  
/lab\_host="EMDH10B"  
/clone\_lib="IGAL - Chicken Intestinal Lymphocyte"  
/note="Organ: Intestine; Vector: PCMV-SPORT6; Site\_1:

Sal; Site 2: NoCI; Normalized library from chicken gut  
infected with coccidia duodenum and middle gut."

Query Match 3.1%; Score 74; DB 5; Length 576;  
Best Local Similarity 100.0%; Pred. No. 2.5e-06;  
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1621 AGAAAGATGGCATTTAAGCTTACCTTGAGAAATAGAGAAATGAGAACTATG 1680  
|||||  
DB 1 AGAAAGATGGCATTTAAGCTTACCTTGAGAAATAGAGAAATGAGAACTATG 60  
QY 1681 ATGGAAGCAATGGG 1694  
|||||  
DB 61 ATGGAAGCAATGGG 74

RESULT 12  
CD733346 480 bp mRNA linear EST 26-JUN-2003  
LOCUS 4045132 IGAL - Chicken Intestinal Lymphocyte Gallus gallus CDNA  
DEFINITION clone IGAL\_41E15 5', mRNA sequence.  
CD733346  
CD733346.1 GI:32284195  
EST.  
SOURCE Gallus gallus (chicken)  
ORGANISM Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
1 (bases 1 to 480)  
Min.W., Lillehoj,H.S., Ashwell,C.M., Matukumalli,L.K., van  
Tassel,C. and Han,J.Y.  
Chicken Intestinal Lymphocyte EST database as a resource for the  
analysis of mucosal immune function  
Unpublished (2003)  
Contact: Hyun S. Lillehoj  
Animal Parasite Diseases Laboratory  
Animal and Natural Resources Institute, USDA  
Bldg.1043, BARC-East, Beltsville, MD 20705, USA  
Tel: 3015048771  
Fax: 3015045103

REFERENCE Email: lillehoj@nri.barc.usda.gov  
AUTHORS Single pass sequencing. Bases called and trimmed with phred  
0.000925 using options -trim\_alt '' -trim\_faasta. Vector identified  
by cross\_match using options -mismatch 12 -minscore 12  
Plate: 41 row: B column: 15  
Seq primer: ATTTAGGTGACACTATAG  
High quality sequence stop: 480.  
Location/Qualifiers  
1..480

FEATURES  
source  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/strain="white leghorn SC"  
/db\_xref="taxon:9031"  
/clone="IGAL\_41E15"  
/sex="mixed"  
/tissue\_type="Gut"  
/cell\_type="Lymphocyte"  
/dev\_stage="Adult"  
/lab\_host="EMDH10B"  
/clone\_lib="IGAL - Chicken Intestinal Lymphocyte"  
/note="Organ: Intestine; Vector: PCMV-SPORT6; Site 1:  
Sal; Site 2: NoCI; Normalized library from chicken gut  
infected with coccidia duodenum and middle gut."

ORIGIN  
Query Match 3.0%; Score 70.4; DB 5; Length 480;  
Best Local Similarity 98.6%; Pred. No. 1.7e-05;  
Matches 71; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1623 AAAGATGGCATTTAAGCTTACCTTGAGAAATAGAGAAATGAGAACTATG 1682  
|||||

Db		745	DDRRBAGTAGKRRRTYKKRWRKRDRDTWDDADBDATDABDRRRGGGCGAAGKGTGRK	686
Oy		2142	AGTTTGATTACAGAGCTGAACA CTGCCTCAGTGAAGAAGTAAGTAA GACTGTGAC	2201
Dd		685	RRRDRATWTRTDAMWADAAMWTTTDTDTDDMDKDRRRRGARRRRTTAAADWWMTWK	626
Oy		2202	TCAGATTCGGGGCTAAGCTCCCTCAACTACAGAAAAAGCACATAAAAATCAAACAT	2261
Dd		625	-AMWMAKMDKTRADRWDRAADDTWTDRARDAWAARAARMRARDBAARADRRWTK	567
Oy		2262	GATGTCATATTTGTTTTCTGCTCGCTGATGATGTATATATATATATATTTTTTTTA	2321
Dd		566	GKTTTATWTTTAAABAAMAWMAWMTTATTTATTTTWTTTTTTTTTTTTTAAWMAA	507
Oy		2322	GGCGTGAATGTGATGAAGAAAGTT	2347
Dd		506	WWTATWAAWTTAAWAAAAAAWMAATT	481
RESULT 14				
BUI123717				
LOCUS			885 bp	mRNA linear EST 25-NOV-2002
DEFINITION			60314739.F1 CSEQCHL18 Gallus gallus cDNA clone CHEST146f24 5', mRNA	
ACCESSION			BUI23717	
VERSION			BUI23717.1	GI:25334366
KEYWORDS			EST.	
SOURCE			Gallus gallus (chicken)	
ORGANISM			Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
			Archaeoptera; Aves; Neognathae; Galliformes; Phasianidae;	
			Phasianinae; Gallus.	
			1 (bases 1 to 885)	
REFERENCE			Boardman, P. E., Sans-Ezquerro, J., Overton, I. M., Burt, D. W., Bosch, E.,	
AUTHORS			Pong, W. T., Tickle, C., Brown, W. R. A., Wilson, S. A. and Hubbard, S. J.	
TITLE			A Comprehensive Collection of Chicken CDNA	
JOURNAL			Curr. Biol. 12 (22), 1965-1969 (2002)	
PUBMED			12445392	
COMMENT			Contact: Simon Hubbard Department of Biomolecular Sciences University of Manchester Institute of Science and Technology (UMIST) PO Box 88, Manchester, M60 1QD, UK Tel: 01612008930 Fax: 01612360409 Email: Simon.Hubbard@umist.ac.uk.	
FEATURES				
source			Location/Qualifiers	
			1..885	
			/organism="Gallus gallus"	
			/mol_type="mRNA"	
			/strain="Compton Line 151"	
			/db_xref="taxon:9031"	
			/clone="CHEST146f24"	
			/sex="Female"	
			/dev_stage="adult"	
			/lab_host="DH10B"	
			/clone_1b="CSEOCHL18"	
			/note="Organ: small intestine; Vector: pBluescript II	
			KS(+); Site_1: EcoRI; Site_2: NotI; Modification of	
			pBluescript II KS(+) [Stratagene] vector to accommodate	
			cDNA produced with the T-primed protocol (Construction of	
			uni-directionally cloned cDNA libraries from messenger RNA	
			for improved 3' end DNA sequencing by Glenn Fv, et al.	
			U.S. Patent # 6,387,624). Cut pbluescript II KS(+) with	
			NotI and EcoRI. Ligate in double stranded adaptor	
			containing BglI and BamHI sites	
			[5'ggcgcgtgacgcccgatccgaaataaag]	
			[5'aattcttttccgatccggcgtagcgc]	
ORIGIN				
Query Match		2.6%	Score 61.4;	DB 3; Length 885;
Best Local Similarity		96.1%;	Pred. No. 0.0027;	

Matches 74; Conservative 0; Mismatches 1; Indels 2; Gaps 1;  
QY 1618 GACGAGAAATGTCATTAAAGTCTTGAAGAAATAGAGAAATGAGAACTATGAAAA 1677  
DB 8 GAGGAGAAAGTGCATTAAAGCTACTTGGAAT--AGAGAAATGAGAACTATGAAAA 65  
QY 1678 TTCATGGAAGCATGGG 1694  
DB 66 TTCATGGAAGCATGGG 82

## RESULT 15

BX276255 443 bp mRNA linear EST 22-FEB-2005  
LOCUS BX276255 AGENAE Gallus gallus multi-tissues normalized library  
DEFINITION (gcag) Gallus gallus cDNA clone gcag0008c.c.24 5prim, mRNA

sequence.

ACCESSION BX276255  
VERSION BX276255.1 GI:28598746

KEYWORDS

SOURCE Gallus gallus (chicken)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

REFERENCE 1 (bases 1 to 443)

AUTHORS

Herault, F., Le Meuth-Metzinger, V., Desert, C., Retout, E., Piumi, F.,  
Klopp, C. and Douaïre, M.

Construction and primary characterization of chicken normalized  
multi-tissue cDNA libraries

Unpublished (2003)

JOURNAL

COMMENT

Contact: Douaïre M  
INRA, UMR INRA-ENSAR Genetique Animale  
65, rue de Saint-Brieuc, RENNES cedex, 35042, FRANCE  
Tel: +33 (0) 2.23.48.54.63  
Fax: +33 (0) 2.23.48.54.70  
Email: Madeleine.Douaïre@roazhon.inra.fr  
Sequence cleaned of vector, adaptor and repetitions. Contact us  
at [agenasupport@jouy.inra.fr](mailto:agenasupport@jouy.inra.fr) to obtain the chromatogram of this  
sequence.  
Plate: 0008 row: c column: 24  
Seq primer: M13R.

FEATURES  
Location/Qualifiers

Source

1..443  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9031"  
/clone="gcag0008c.c.24"  
/tissue\_type="multi-tissues"  
/dev\_stage="from embryos to adults"  
/lab\_host="DH10B"  
/clone\_lib="AGENAE Gallus gallus multi-tissues normalized  
library (gcag)"  
/note="Vector: pT7T3D-PacI; tissue: adipose tissue, brain,  
intestine, liver, multi-tissues, muscle, ovary,  
testis, adrenal gland, bone marrow, caecum, duodenum,  
embryos, fabricius gland, genital tract, granulosa,  
hypothalamus, ileon, jejunum, lymphoid tissue, oviduct,  
pancreas, skin, spleen, thymus, utero-vaginal gland,  
pituitary gland, small follicle, hematopoietic progenitor  
cells. Clone distribution: AGENAE Resource centre.  
Francis Piumi, Francis.Piumi@jouy.inra.fr, INRA, CEA  
Radiobiologie et Etude du genome (UREG), Domaine de  
Vilvert, 78352, Jouy-en-Josas cedex, FRANCE, +33 (0)  
1.34.65.28.02, +33 (0) 1.34.65.22.73"

## ORIGIN

Query Match 2 5%; Score 60.4; DB 4; Length 443;

Best Local Similarity 98.4%; Pred. No. 0.004; 1; Indels 0; Gaps 0;  
Matches 61; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2320 TAGGCGTAATGTGATGAAAGAAAGTTAGAGCCCAAGATATCTGAAGCTCACTATTC 2379

DB 31 TGCGCGTAATGTGATGAAAGAAAGTTAGAGCCCAAGATATCTGAAGCTCACTATTC 90

QY 2380 AG 2381

DB 91 AG 92

Search completed: May 28, 2006, 14:46:40  
Job time : 12210.9 secs







```
/ FILE REFERENCE: CCDNA-260XX
/ CURRENT APPLICATION NUMBER: US/09/385,982
/ CURRENT FILING DATE: 1999-08-30
/ EARLIER APPLICATION NUMBER: 09/328,111
/ EARLIER FILING DATE: 1999-06-08
/ EARLIER APPLICATION NUMBER: 60/117,393
/ EARLIER FILING DATE: 1999-01-27
/ EARLIER APPLICATION NUMBER: 60/098,639
/ EARLIER FILING DATE: 1998-08-31
/ NUMBER OF SEQ ID NOS: 544
/ SOFTWARE: FASTSEQ for Windows Version 3.0
/ SEQ ID NO 474
/ LENGTH: 515
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)-(515)
/ OTHER INFORMATION: n = A,T,C or G
US-09-385-982-474
```

```
Query Match
Best Local Similarity 1.9%; Score 44.2; DB 3; Length 515;
Matches 55; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
```

```
QY 1623 AAAGATGGCATTTAAGCGTACTTGAAATAAGAGAAAAATGAGAACTATGAAAAATTCAT 1682
DB 23 AATCATGGCGTTTGACAGCAGCTTGAAGGTAGACCGAGTGAATACTATGACAAGTTCA 82
QY 1683 GGAAGCATGGGT 1695
DB 83 GGAATAAATGGT 95
```

```
RESULT 7
US-09-949-016-4519
/ Sequence 4519, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CLO01307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 4519
/ LENGTH: 2252
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-4519
```

```
Query Match
Best Local Similarity 1.9%; Score 44.2; DB 3; Length 2252;
Matches 55; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
```

```
QY 1623 AAAGATGGCATTTAAGCGTACTTGAAATAAGAGAAAAATGAGAACTATGAAAAATTCAT 1682
DB 58 AATCATGGCGTTTGACAGCAGCTTGAAGGTAGACCGAGTGAATACTATGACAAGTTCA 117
QY 1683 GGAAGCATGGGT 1695
DB 118 GGAATAAATGGT 130
```

RESULT 8

```
US-09-949-016-19
/ Sequence 19, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CLO01307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 19
/ LENGTH: 2273
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-19
```

```
Query Match
Best Local Similarity 1.9%; Score 44.2; DB 3; Length 2273;
Matches 55; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
```

```
QY 1623 AAAGATGGCATTTAAGCGTACTTGAAATAAGAGAAAAATGAGAACTATGAAAAATTCAT 1682
DB 58 AATCATGGCGTTTGACAGCAGCTTGAAGGTAGACCGAGTGAATACTATGACAAGTTCA 117
QY 1683 GGAAGCATGGGT 1695
DB 118 GGAATAAATGGT 130
```

```
RESULT 9
US-08-232-463-14/C
/ Sequence 14, Application US/08232463
/ Patent No. 5670367
/ GENERAL INFORMATION:
/ APPLICANT: DOMNER, F.
/ APPLICANT: SCHEIFLINGER, F.
/ APPLICANT: FALKNER, F. G.
/ TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
/ NUMBER OF SEQUENCES: 52
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Foley & Lardner
/ STREET: 1800 Diagonal Road, Suite 500
/ CITY: Alexandria
/ STATE: VA
/ COUNTRY: USA
/ ZIP: 22313-0299
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/232,463
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/07/935,313
/ FILING DATE:
/ APPLICATION NUMBER: EP 91 114 300.6
/ FILING DATE: 26-AUG-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BENT, Stephen A.
/ REGISTRATION NUMBER: 29,768
/ REFERENCE/DOCKET NUMBER: 30472/114 IMMU
/ TELECOMMUNICATION INFORMATION:
```



TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: pTZ9pt-F18  
US-08-232-463-14

Query Match 1.8%; Score 43.8; DB 2; Length 7218;  
Best Local Similarity 8.4%; Pred. No. 0.49; Mismatches 191; Indels 0; Gaps 0;  
Matches 36; Conservative 204; Mismatches 191; Indels 0; Gaps 0;

Qy 628 TATCTGTGTAAACAATTCTTTCAGCATTCAGATTACATTTTGAAGTAATA 687  
Db 1474 TATCTATGCAAGTAGTTAAAGATAGAAATTTGTACRRRRRRRRRRRRRRR 1415  
Qy 688 GACGACGATTTGGTCCGTCATAGAAACAGACTAATATATCTGAGTTAGTACA 747  
Db 1414 RR 1355  
Qy 748 GCAGATTAGCACAGCAAAATTTGCTCAGTTCAAGTACATCTTGGGGAAGAAG 807  
Db 1354 RR 1295  
Qy 808 GAGCTGAGCCAGTGTGCTCATTTTCTGATTAATCTTCAACATTTAAACCTGGATC 867  
Db 1294 RR 1235  
Qy 868 TATGGAATCAAAACGTTGGTAAATTCATTAAGCAGACATCACTAGTAGAAT 927  
Db 1234 RR 1175  
Qy 928 GGACAGAAACAGACATTCAGTAATGGCTATAATATAGAAATAGTAGAGTGTCC 987  
Db 1174 RR 1115  
Qy 988 TGAATTAGACTTATTAAGAAGTAGAGACAGAAATGAAATATCGCAATTTCT 1047  
Db 1114 RR 1055  
Qy 1048 GTAGCTAGCA 1058  
Db 1054 CTCGACCTGCA 1044

RESULT 10  
US-09-621-976-2813  
Sequence 2813, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Mline Edwards, J.B.  
APPLICANT: Jobert, S.Y.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.ppm  
SEQ ID NO 2813  
LENGTH: 832  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 235..399  
US-09-621-976-2813

Query Match 1.8%; Score 43; DB 3; Length 832;  
Best Local Similarity 15.0%; Pred. No. 0.31;  
Matches 51; Conservative 147; Mismatches 142; Indels 1; Gaps 1;

Qy 1073 TTCTGAAACGACCGAGTTTCCAAACTACTGATGATGTTCACTGATGATCTTCATCTC 1132  
Db 35 KTYWRMRKKKAAWKKWTWYWRVYMWGTYKKKMKCRKTKYKKKKKKYMMWYWG 94  
Qy 1133 ATGCTTATATGAGAGATAGATGATCTCAACCAATATTAAGTAAAGCAAGACAGAT 1192  
Db 95 RSTYMAAMTRITWGYATYRSMMYRRICMKKAYIRRTTTCSSSGWYMKRKAATW 154  
Qy 1193 TGTGTTT-ATCTGTGGTAATAACGTTTCTCAGTGTATAAAGACCTCCACGAG 1251  
Db 155 MMKXTYYAATRYMMWMMCTKRPASWYCWMMGARSKMSTRSRSYASARSARCCYS 214  
Qy 1252 TATAAGCTTATGCAACAAAGAAATGTCATATCTCTTACTCTCATTTATTTT 1311  
Db 215 CSMGMSWKYMRMRMGATGAGMKMRASCMRRRYAGSKTSYSMMWCWTRSMKY 274  
Qy 1312 CATTAGATAGCCGGTTTCTCAACTCAATTAAGTAAAGACAGATGAGTGGTACT 1371  
Db 275 CTTAKRTGYTCYKGGWGRGRYASKTMMKMMWCMAMRYSTGTASMMWRMY 334  
Qy 1372 GACTGTTATAAGAGAGATATAAGATCTATCATCTATT 1412  
Db 335 YTMAMKWKYAMADBAWRMMWMMAMRRACAAATATATATT 375

RESULT 11  
US-09-949-016-18019/c  
Sequence 18019, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CLO01307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 18019  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-18019

Query Match 1.8%; Score 42.6; DB 3; Length 601;  
Best Local Similarity 54.0%; Pred. No. 0.34; Mismatches 74; Indels 0; Gaps 0;  
Matches 87; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Qy 2218 AGCTCCCTCACTACAGAAAAGTCACATTAATAAATGCAACATGATCTTATTTGTT 2277  
Db 258 AGTTTCTCCCACTCAAAAATTAACAAACAAACATTAACCTTGAAAAATTAACACTTC 199  
Qy 2278 TTCTCTGCTGAGTGAATATGATTAATATTTTGGGGGTGAATGATGA 2337  
Db 198 CTATGAGATTGACTTATTTTCTCAATGCTTACCTTTACAGGTGTTAATAGTGA 139  
Qy 2338 AAGAAAGTTAGAGCCACAGATATCTGAAGCTCAGTATT 2378  
Db 138 AAGAAAGCTTGAGCTCAGTCAATTTGAAGCTGAGCAATT 98

RESULT 12  
US-09-949-016-18020/c

```
; Sequence 18020, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 18020
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-18020
```

```
Query Match
Best Local Similarity 1.8%; Score 42.6; DB 3; Length 601;
Matches 87; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
```

```
QY 2218 AGCTCCCTCACTACAGAAAAGTCACATTAATAATGCAACATGATGTTCTATTGTT 2277
DB 554 AGTTTCTTCCCACTCCCTCAAAAATTAACAAACAAACATACCTTGAAAAATTAACACTTC 495
QY 2278 TTCTCTGCTGATGATTAATGATTAATTAATTTTATTTTATTTTATTTTATTTTATTTT 2337
DB 494 CTATGGATTTGACTTATTTATTTCTCATTTGCTTCACTTTTACAGGTATTAATATGTA 435
QY 2338 AAAAGAAAGTTAGAGCCCAAGATTAATCTGAAGCTCACTATT 2378
DB 434 AAAAGAAAGTTGACGCTCATGACATTTGAAGCTGACAAATT 394
```

```
RESULT 13
US-09-949-016-161279/c
; Sequence 161279, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 161279
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-161279
```

```
Query Match
Best Local Similarity 1.8%; Score 42.6; DB 3; Length 601;
Matches 87; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
```

```
QY 2218 AGCTCCCTCACTACAGAAAAGTCACATTAATAATGCAACATGATGTTCTATTGTT 2277
DB 258 AGTTTCTTCCCACTCCCTCAAAAATTAACAAACAAACATACCTTGAAAAATTAACACTTC 199
```

```
QY 2278 TTCTCTGCTGATGATTAATGATTAATTAATTTTATTTTATTTTATTTTATTTTATTTT 2337
DB 198 CTATGGATTTGACTTATTTATTTCTCATTTGCTTCACTTTTACAGGTATTAATATGTA 139
QY 2338 AAAAGAAAGTTAGAGCCCAAGATTAATCTGAAGCTCACTATT 2378
DB 138 AAAAGAAAGTTGACGCTCATGACATTTGAAGCTGACAAATT 98
```

```
RESULT 14
US-09-949-016-161280/c
; Sequence 161280, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 161280
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-161280
```

```
Query Match
Best Local Similarity 1.8%; Score 42.6; DB 3; Length 601;
Matches 87; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
```

```
QY 2218 AGCTCCCTCACTACAGAAAAGTCACATTAATAATGCAACATGATGTTCTATTGTT 2277
DB 554 AGTTTCTTCCCACTCCCTCAAAAATTAACAAACAAACATACCTTGAAAAATTAACACTTC 495
QY 2278 TTCTCTGCTGATGATTAATGATTAATTAATTTTATTTTATTTTATTTTATTTTATTTT 2337
DB 494 CTATGGATTTGACTTATTTATTTCTCATTTGCTTCACTTTTACAGGTATTAATATGTA 435
QY 2338 AAAAGAAAGTTAGAGCCCAAGATTAATCTGAAGCTCACTATT 2378
DB 434 AAAAGAAAGTTGACGCTCATGACATTTGAAGCTGACAAATT 394
```

```
RESULT 15
US-09-385-982-376/c
; Sequence 376, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CDDNA-260XX
; CURRENT FILING DATE: 1999-08-30
; PRIOR APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 376
; LENGTH: 611
; TYPE: DNA
```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(611)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-376

```

```

Query Match          1.8%; Score 42.6; DB 3; Length 611;
Best Local Similarity 55.0%; Pred. No. 0.35;
Matches 72; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

```

```

QY 2220 CTCCTCACTACAGAAAAAGTCACATRAAAATGCAACATGATGTTCTATTGTTT 2279
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 351 CNCNTTATATTCNGTTTACGAAACNAAAAATGCGNACNTTTTTTTTTTTTTT 292
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 2280 TCTCGCTTGATGTTAATTGATTATTATTATTTTTTTTTTTAGCGGTGATGTGAA 2339
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 291 TTTTTTTTGNTTTTTTTTTTTTTTTTTTTTTTTTTTTATAGATTATTATTAATA 232
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 2340 AGAAAGTTAG 2350
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 231 GGGGAGATAG 221
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: May 27, 2006, 19:25:36  
 Job time : 477.706 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

# OM nucleic - nucleic search, using sw model

Run on: May 27, 2006, 18:56:27 ; Search time 3252.58 Seconds  
(without alignments)  
8994.972 Million cell updates/sec

Title: US-10-099-663-1  
Perfect score: 2381  
Sequence: 1 agcttcctgcgcagaaag.....atcgaagctcattcag 2381

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

## Database :

Published Applications\_MA.Main:\*  
1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
7: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
8: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
9: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*  
10: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*  
11: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*  
12: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10G\_PUBCOMB.seq:\*  
13: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11\_PUBCOMB.seq:\*  
14: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11B\_PUBCOMB.seq:\*  
15: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11C\_PUBCOMB.seq:\*  
16: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11D\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2381	100.0	2381	US-10-099-663-1	Sequence 1, Appli
2	336	14.1	336	US-10-099-663-2	Sequence 2, Appli
3	50.6	2.1	16914	US-10-741-601-5698	Sequence 5, Appli
4	50.6	2.1	16914	US-10-741-600-17777	Sequence 17777, A
5	50.6	2.1	16914	US-10-995-561-13349	Sequence 13349, A
6	50	2.1	564	US-11-136-527-3784	Sequence 3784, Ap
7	50	2.1	564	US-11-136-527-7880	Sequence 7880, Ap
8	49.8	2.1	630	US-10-027-632-5694	Sequence 5694, Ap
9	49.8	2.1	630	US-10-027-632-5694	Sequence 5694, Ap
10	49.8	2.1	630	US-10-116-275-353	Sequence 303, App
11	49.2	2.1	426	US-11-085-792-3	Sequence 3, Appli
12	48	2.0	546	US-09-925-065A-51803	Sequence 51803, A
13	48	2.0	546	US-09-925-065A-51803	Sequence 51803, A
14	48	2.0	546	US-10-301-480-153041	Sequence 153041, A
15	48	2.0	546	US-10-301-480-153041	Sequence 153041, A
16	48	2.0	568	US-10-301-480-288831	Sequence 288831, App
17	48	2.0	568	US-10-301-480-288831	Sequence 288831, App

C 18	48	2.0	569	4	US-09-925-065A-200196	Sequence 200196,
C 19	48	2.0	569	5	US-09-925-065A-200196	Sequence 200196,
C 20	48	2.0	570	4	US-09-925-065A-200197	Sequence 200197,
C 21	48	2.0	570	5	US-09-925-065A-200197	Sequence 200197,
C 22	48	2.0	599	12	US-10-301-480-288830	Sequence 288830,
C 23	48	2.0	599	12	US-10-301-480-288830	Sequence 288830,
C 24	48	2.0	814	4	US-09-925-065A-63272	Sequence 63272, A
C 25	48	2.0	814	5	US-09-925-065A-63272	Sequence 63272, A
C 26	48	2.0	814	12	US-10-301-480-164510	Sequence 164510,
C 27	48	2.0	814	12	US-10-301-480-164510	Sequence 164510,
C 28	48	2.0	991	12	US-10-301-480-584425	Sequence 584425,
C 29	48	2.0	991	12	US-10-301-480-1197834	Sequence 1197834,
C 30	48	2.0	992	12	US-10-301-480-584426	Sequence 584426,
C 31	48	2.0	992	12	US-10-301-480-1197835	Sequence 1197835,
C 32	47.6	2.0	399	13	US-11-085-792-1	Sequence 1, Appli
C 33	47.2	2.0	337	3	US-09-960-352-6036	Sequence 6036, Ap
C 34	47	2.0	2636	13	US-11-097-143-21061	Sequence 21061, A
C 35	46.8	2.0	662	9	US-10-335-053-36	Sequence 36, Appli
C 36	46.8	2.0	670	10	US-10-764-420-2245	Sequence 2245, Ap
C 37	46.2	1.9	6222	7	US-10-311-455-665	Sequence 665, App
C 38	44.4	1.9	565	4	US-09-925-065A-110075	Sequence 310075,
C 39	44.4	1.9	565	5	US-09-925-065A-110075	Sequence 310075,
C 40	44.4	1.9	573	12	US-10-301-480-385620	Sequence 385620,
C 41	44.4	1.9	573	12	US-10-301-480-385620	Sequence 385620,
C 42	44.2	1.9	515	3	US-09-871-161-474	Sequence 474, App
C 43	44.2	1.9	2252	8	US-10-741-601-261	Sequence 261, App
C 44	44.2	1.9	2252	9	US-10-741-600-745	Sequence 745, App
C 45	44.2	1.9	2252	10	US-10-995-561-449	Sequence 449, App

## ALIGNMENTS

RESULT 1	US-10-099-663-1	Application US/10099663
Sequence 1, Application US/10099663	Publication No. US20030177516A1	GENERAL INFORMATION:
APPLICANT: Avigence, Inc	TITLE OF INVENTION: Avian Gut-Specific Promoters	FILE REFERENCE: A181
CURRENT FILING DATE: 2002-03-14	NUMBER OF SEQ ID NOS: 19	SOFTWARE: PatentIn version 3.0
SEQ ID NO 1	LENGTH: 2381	TYPE: DNA
ORGANISM: Gallus gallus	FEATURES:	NAME/KEY: 5'UTR
LOCATION: (1)..(1626)	NAME/KEY: exon	LOCATION: (1627)..(1693)
NAME/KEY: Intron	LOCATION: (1694)..(2322)	NAME/KEY: exon
LOCATION: (2333)..(2381)	NAME/KEY: Intron	LOCATION: (2381)
Query Match	100.0%; Score 2381; DB 7; Length 2381;	Best Local Similarity 100.0%; Pred. No. 0;
Matches 2381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 AGCTTCTGCGCAGAAAGGCTGGGCTTCTGTTCCCTACACAGCTTAAGCAATC	60
DB	1 AGCTTCTGCGCAGAAAGGCTGGGCTTCTGTTCCCTACACAGCTTAAGCAATC	60
QY	61 CCCAAGTCAAAAGCGGCTGTAAGAGAGATGCTCACTTCAATGAATGATATG	120
DB	61 CCCAAGTCAAAAGCGGCTGTAAGAGAGATGCTCACTTCAATGAATGATATG	120
QY	121 AATATATCATAAACGAGCTCTGTGGCAGATCAGAGATTAACCTCTGCGACAAAT	180

```
Db 121 AAATATCATAAAAGAGCTCTGTTGGCAGATCAGAGTAACTCTCTGGAGCAAAATTT 180
Qy 181 CTTAAAGGTATAGGTAGAACAGAGAGTGTTCGACCTAAATGGAATGATTTGACACA 240
Db 181 CTTAAAGGTATAGGTAGAACAGAGAGTGTTCGACCTAAATGGAATGATTTGACACA 240
Qy 241 TTTGATCTTCTAGGAGACAAAAGGTCGTGAAAACAATTAATTCTGTGTGACAGTCAGT 300
Db 241 TTTGATCTTCTAGGAGACAAAAGGTCGTGAAAACAATTAATTCTGTGTGACAGTCAGT 300
Qy 301 AGCAGCTGTTTTGGGTGCACTACAGCACTTTGTTGCAACATTAACAATCTAAGTTG 360
Db 301 AGCAGCTGTTTTGGGTGCACTACAGCACTTTGTTGCAACATTAACAATCTAAGTTG 360
Qy 361 TTTTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
Db 361 TTTTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
Qy 421 TACTTCCCTCTGCATCTCTTACCCAGATTAGCATTGATTTCAAAATGAACCTGAGTGA 480
Db 421 TACTTCCCTCTGCATCTCTTACCCAGATTAGCATTGATTTCAAAATGAACCTGAGTGA 480
Qy 481 ATGGAAGCCACATATTTTGTATGACAGCAAAAGTTCTAAATTTATAGTTATCTTCA 540
Db 481 ATGGAAGCCACATATTTTGTATGACAGCAAAAGTTCTAAATTTATAGTTATCTTCA 540
Qy 541 GTAAAACTTTTGTCTGAGGTCTGAGAGAAAAAGAAATTTATGATTAACAGCAGTACGA 600
Db 541 GTAAAACTTTTGTCTGAGGTCTGAGAGAAAAAGAAATTTATGATTAACAGCAGTACGA 600
Qy 601 AAATCATTAAGTAAAGGCCAACCCCTGTTATCTGTGTGATAGCAACATTCAATTCAGCAT 660
Db 601 AAATCATTAAGTAAAGGCCAACCCCTGTTATCTGTGTGATAGCAACATTCAATTCAGCAT 660
Qy 661 TCAGATTTTACATTTTGTAGAGCTAAATAGACAGACATTTGGGCCCTCATAGGAACGA 720
Db 661 TCAGATTTTACATTTTGTAGAGCTAAATAGACAGACATTTGGGCCCTCATAGGAACGA 720
Qy 721 CTAACTATTAATCCCTAGTTTATGATCAAGCAATTTTGAACAAGCAAAATTTGCTCAGTTTC 780
Db 721 CTAACTATTAATCCCTAGTTTATGATCAAGCAATTTTGAACAAGCAAAATTTGCTCAGTTTC 780
Qy 781 AAGTACATATCTTGTGGGAGAGAGAGAGGCGCAGTGTGTGCTCATTTTCTGCAAT 840
Db 781 AAGTACATATCTTGTGGGAGAGAGAGAGGCGCAGTGTGTGCTCATTTTCTGCAAT 840
Qy 841 ATCTTCAACATTTTAAACCTGGGATCTATGGAATCAAAACGTTGGGTAAATTCAT 900
Db 841 ATCTTCAACATTTTAAACCTGGGATCTATGGAATCAAAACGTTGGGTAAATTCAT 900
Qy 901 TAGCAGCATCACTACTGTAGGAATGAGACAGAAAGAGCATTCCTGATAGGGCTAT 960
Db 901 TAGCAGCATCACTACTGTAGGAATGAGACAGAAAGAGCATTCCTGATAGGGCTAT 960
Qy 961 AATTAAGAAATAGAGAGAGTCTCTGAATTTTGAATCTAATTTAAAGTGAAGACA 1020
Db 961 AATTAAGAAATAGAGAGAGTCTCTGAATTTTGAATCTAATTTAAAGTGAAGACA 1020
Qy 1021 CGAATGAGAAATATCAATGCAATTTCTGTAGCTCAGACATAGCATGAAAGTGTCTGAAA 1080
Db 1021 CGAATGAGAAATATCAATGCAATTTCTGTAGCTCAGACATAGCATGAAAGTGTCTGAAA 1080
Qy 1081 CTGAACCGAGTTTCCCAAATCTACTGTGATGTTCAAGTGAATCTTCAATCTGATTTAT 1140
Db 1081 CTGAACCGAGTTTCCCAAATCTACTGTGATGTTCAAGTGAATCTTCAATCTGATTTAT 1140
Qy 1141 TAATGAGAGTAAATAGATTTCTCAACAATTTAGAAATGACAAAGCAGAGATTTGTTTT 1200
Db 1141 TAATGAGAGTAAATAGATTTCTCAACAATTTAGAAATGACAAAGCAGAGATTTGTTTT 1200
Qy 1201 ATCTGTGGGTAAATAGTGTCTTCCAGTTGTTAAAGACCTCCCAACAGATTAAGTC 1260
Db 1201 ATCTGTGGGTAAATAGTGTCTTCCAGTTGTTAAAGACCTCCCAACAGATTAAGTC 1260
Db 1201 ATCTGTGGGTAAATAGTGTCTTCCAGTTGTTAAAGACCTCCCAACAGATTAAGTC 1260
Qy 1261 CTATGCAACAAAGAAATAGTCAATTAATCTCTTAAGTCTCATTTATTTTCATTAAGATA 1320
Db 1261 CTATGCAACAAAGAAATAGTCAATTAATCTCTTAAGTCTCATTTATTTTCATTAAGATA 1320
Qy 1321 GCCGGTTTTTACTACAACTTCAATTAAGATGAACAGATATGAGTGTAGCTGTTTAT 1380
Db 1321 GCCGGTTTTTACTACAACTTCAATTAAGATGAACAGATATGAGTGTAGCTGTTTAT 1380
Qy 1381 TAAAGAGATTAATTAAGATTAATCTATCATTTTGAAGCAATTAAGGAGGAGATTCAG 1440
Db 1381 TAAAGAGATTAATTAAGATTAATCTATCATTTTGAAGCAATTAAGGAGGAGATTCAG 1440
Qy 1441 CAAACAGTGTCTTACAAAGTGAAGAAACAAGTTAACTTAAAGTGAACCCCTCTGACAA 1500
Db 1441 CAAACAGTGTCTTACAAAGTGAAGAAACAAGTTAACTTAAAGTGAACCCCTCTGACAA 1500
Qy 1501 GATCAATGCCACAGTTGAGCTTTAGCCAGCCACATCATATGTAATTTGCTTCCGATTA 1560
Db 1501 GATCAATGCCACAGTTGAGCTTTAGCCAGCCACATCATATGTAATTTGCTTCCGATTA 1560
Qy 1561 AGCCTGTTCAATAATTTCTTTGCAAGCTCTGCTACTTACAGAAAGTGTGCTTACAGAC 1620
Db 1561 AGCCTGTTCAATAATTTCTTTGCAAGCTCTGCTACTTACAGAAAGTGTGCTTACAGAC 1620
Qy 1621 AGAAGATGAGCATTTAAGGGTATCTTGAGAAATAGAGAAATAGAGAACTATGAAAAATTC 1680
Db 1621 AGAAGATGAGCATTTAAGGGTATCTTGAGAAATAGAGAAATAGAGAACTATGAAAAATTC 1680
Qy 1681 ATGGAAGCATAGGAGTAAAGCTTACTTTTGAATGCTTTTGAAGCAAGATCAACATAC 1740
Db 1681 ATGGAAGCATAGGAGTAAAGCTTACTTTTGAATGCTTTTGAAGCAAGATCAACATAC 1740
Qy 1741 GCGGGAATCAAACTTAAGCTGTTCATGAACTAATCCATGCTGCTCTTGTGTGT 1800
Db 1741 GCGGGAATCAAACTTAAGCTGTTCATGAACTAATCCATGCTGCTCTTGTGTGT 1800
Qy 1801 CTGCTATTTTGGCCTTGAACATTTGCTGCACTTATTTTGAAGAACTCTATAGAGGGA 1860
Db 1801 CTGCTATTTTGGCCTTGAACATTTGCTGCACTTATTTTGAAGAACTCTATAGAGGGA 1860
Qy 1861 ATACAAGAGAGAAAAACATTCGATTTTATTTGCAATGCGAATCTTATAGCATTTAGCT 1920
Db 1861 ATACAAGAGAGAAAAACATTCGATTTTATTTGCAATGCGAATCTTATAGCATTTAGCT 1920
Qy 1921 AATTCAGATTAAGAGGATTCGACGAGAAATTTAAATAGAAATTAATGTAAGAAATTAAT 1980
Db 1921 AATTCAGATTAAGAGGATTCGACGAGAAATTTAAATAGAAATTAATGTAAGAAATTAAT 1980
Qy 1981 TTGATTAAGATCTGTTGAAAAATTAACAAGAGGGAATTTGCTGCTCCAGTTTTCAGA 2040
Db 1981 TTGATTAAGATCTGTTGAAAAATTAACAAGAGGGAATTTGCTGCTCCAGTTTTCAGA 2040
Qy 2041 ACAACATGATTTAGTCAATTTTAAACATGCTAGTCTTACTTTAAGCTGTGACAACTGC 2100
Db 2041 ACAACATGATTTAGTCAATTTTAAACATGCTAGTCTTACTTTAAGCTGTGACAACTGC 2100
Qy 2101 CTGTAAATGAGATTAACAATTAATCTAATCCATGTTGATAGTATGATTTGATTAACAGCTGA 2160
Db 2101 CTGTAAATGAGATTAACAATTAATCTAATCCATGTTGATAGTATGATTTGATTAACAGCTGA 2160
Qy 2161 ACAGTGCCTCAGTGAAGGTGAGAGAGATTAAGTCTGAGTCAAAATCTGGGCTTAAGC 2220
Db 2161 ACAGTGCCTCAGTGAAGGTGAGAGAGATTAAGTCTGAGTCAAAATCTGGGCTTAAGC 2220
Qy 2221 TCCCTCAACTACAGAAAAAGTCAATTAATAATGCAAAATGATATGTTGTTTT 2280
Db 2221 TCCCTCAACTACAGAAAAAGTCAATTAATAATGCAAAATGATATGTTGTTTT 2280
Qy 2281 CTCTGCTGATGTTAATTAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 2340
Db 2281 CTCTGCTGATGTTAATTAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 2340
Qy 2340 CTCTGCTGATGTTAATTAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 2340
Db 2340 CTCTGCTGATGTTAATTAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 2340
```

Qy 2341 GAAAGTTAGAGCCACGATATCTGAAGCTCACTATTCAAG 2381  
 Db 2241 GAAAGTTAGAGCCACGATATCTGAAGCTCACTATTCAAG 2381

RESULT 2

US-10-099-663-2  
 ; Sequence 2, Application US/10099663  
 ; Publication No. US20030177516A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Avigenics, Inc  
 ; TITLE OF INVENTION: Avian GUT-Specific Promoters  
 ; FILE REFERENCE: A181  
 ; CURRENT APPLICATION NUMBER: US/10/099,663  
 ; CURRENT FILING DATE: 2002-03-14  
 ; NUMBER OF SEQ ID NOS: 19  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 2  
 ; LENGTH: 336  
 ; TYPE: DNA  
 ; ORGANISM: Gallus gallus  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)..(336)  
 ; OTHER INFORMATION: chick intestinal fatty acid binding protein promoter region  
 US-10-099-663-2

Query Match 14.1%; Score 336; DB 7; Length 336;  
 Best Local Similarity 100.0%; Pred. No. 8.6e-74;  
 Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1301 ATATATATTTTTCATTAGATAGCCGGTTTTTTTACTACAACTCAATTAATAGATGACAGAAATG 1360  
 Db 1 ATATATATTTTTCATTAGATAGCCGGTTTTTTTACTACAACTCAATTAATAGATGACAGAAATG 60  
 Qy 1361 AATGGTTAGTACTGTTTATTAAGAGATTAATTAAGATCTATCATCTATTGAGGCA 1420  
 Db 61 AATGGTTAGTACTGTTTATTAAGAGATTAATTAAGATCTATCATCTATTGAGGCA 120  
 Qy 1421 TAAGGGAGGAGAGATTCAGCAACAGTGTCTTCAAGTGGAAAAACAAGTTAACTTAA 1480  
 Db 121 TAAGGGAGGAGAGATTCAGCAACAGTGTCTTCAAGTGGAAAAACAAGTTAACTTAA 180  
 Qy 1481 GTGACCCCTCTCTTGCAGCAATGCAAGTGGAGCTTTTGGCCAGCAATCATCA 1540  
 Db 181 GTGACCCCTCTCTTGCAGCAATGCAAGTGGAGCTTTTGGCCAGCAATCATCA 240  
 Qy 1541 TGTAAATGCTTTCCTGATTAAGCCTGTTCAATTAATCTCTTGGAAAGCTCTGCTACTTA 1600  
 Db 241 TGTAAATGCTTTCCTGATTAAGCCTGTTCAATTAATCTCTTGGAAAGCTCTGCTACTTA 300  
 Qy 1601 CCAGAACTGCTGCTACAGACAGAAAGATGGCATTTA 1636  
 Db 301 CCAGAACTGCTGCTACAGACAGAAAGATGGCATTTA 336

RESULT 3

US-10-741-601-5698  
 ; Sequence 5698, Application US/10741601  
 ; Publication No. US20040166519A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CARGILL, Michele et al.  
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
 ; FILE REFERENCE: C1001500  
 ; CURRENT APPLICATION NUMBER: US/10/741,601  
 ; CURRENT FILING DATE: 2003-12-22  
 ; NUMBER OF SEQ ID NOS: 26415  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 5698  
 ; LENGTH: 16914  
 ; TYPE: DNA

; ORGANISM: Homo sapiens  
 US-10-741-601-5698

Query Match 2.1%; Score 50.6; DB 8; Length 16914;  
 Best Local Similarity 58.8%; Pred. No. 0.69;  
 Matches 124; Conservative 2; Mismatches 76; Indels 9; Gaps 2;

Qy 1516 TGAGCTTTAGCCACGACATCATGTAATTCCTTCTGATTAAGCTCTTCAATTAAT 1575  
 Db 5945 TGAACCTTTAGCTTCCACATCATGTAATTCCTTCTGATTAAGCTCTTCAATTAAT 6004  
 Qy 1576 TCTC-----TTTGCAAGCTCTGCTACTTACCAAGTCTGCTACAGCAAGAA 1627  
 Db 6005 TCTGCCCCAAGACAGACCTGAATCTTACGCTGCTGAGGCTGACT-CAACTGAAATCA 6063  
 Qy 1628 TGGCATTTTAACGGTACTTGGAAAAATAGAAAAATGAACTATGAAAAATTCATGGAAG 1687  
 Db 6064 TGGCGTTTGACAGCATTTGGAAAGGTAGCCGAGGTGAAACTATGACAAAGTTCAATGAAA 6123  
 Qy 1688 CAATGGGTAAAGCTTAACTTTTGAATGCTT 1718  
 Db 6124 AAATGGGTAAAGACTTTATTTCTTTGTGGCT 6154

RESULT 4

US-10-741-600-17777  
 ; Sequence 17777, Application US/10741600  
 ; Publication No. US20050026169A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CARGILL, Michele et al.  
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
 ; FILE REFERENCE: C1001499  
 ; CURRENT APPLICATION NUMBER: US/10/741,600  
 ; CURRENT FILING DATE: 2003-12-22  
 ; NUMBER OF SEQ ID NOS: 7397  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 17777  
 ; LENGTH: 16914  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-741-600-17777

Query Match 2.1%; Score 50.6; DB 9; Length 16914;  
 Best Local Similarity 58.8%; Pred. No. 0.69;  
 Matches 124; Conservative 2; Mismatches 76; Indels 9; Gaps 2;

Qy 1516 TGAGCTTTAGCCACGACATCATGTAATTCCTTCTGATTAAGCTCTTCAATTAAT 1575  
 Db 5945 TGAACCTTTAGCTTCCACATCATGTAATTCCTTCTGATTAAGCTCTTCAATTAAT 6004  
 Qy 1576 TCTC-----TTTGCAAGCTCTGCTACTTACCAAGTCTGCTACAGCAAGAA 1627  
 Db 6005 TCTGCCCCAAGACAGACCTGAATCTTACGCTGCTGAGGCTGACT-CAACTGAAATCA 6063  
 Qy 1628 TGGCATTTTAACGGTACTTGGAAAAATAGAAAAATGAACTATGAAAAATTCATGGAAG 1687  
 Db 6064 TGGCGTTTGACAGCATTTGGAAAGGTAGCCGAGGTGAAACTATGACAAAGTTCAATGAAA 6123  
 Qy 1688 CAATGGGTAAAGCTTAACTTTTGAATGCTT 1718  
 Db 6124 AAATGGGTAAAGACTTTATTTCTTTGTGGCT 6154

RESULT 5

US-10-995-561-13349  
 ; Sequence 13349, Application US/10995561  
 ; Publication No. US20050272054A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CARGILL, Michele et al.  
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
 ; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
 ; TITLE OF INVENTION: DETECTION AND USES THEREOF

FILE REFERENCE: CU001559  
CURRENT APPLICATION NUMBER: US/10/995,561  
CURRENT FILING DATE: 2004-11-24  
NUMBER OF SEQ ID NOS: 85702  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13349  
LENGTH: 16914  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-995-561-13349

Query Match 2.1%; Score 50.6; DB 10; Length 16914;  
Best Local Similarity 58.8%; Pred. No. 0.69;  
Matches 124; Conservative 2; Mismatches 76; Indels 9; Gaps 2;

QY 1516 TGACCTTTAGCCACCATCATCATTAATTTGCTTCCGTATGAGAGCTGTCTATTAAT 1575  
DB 5945 TGAACCTTTAGCTTCCACATCACTATGAAAGTTGTTAAGATTAAGAAATATTAAT 6004  
QY 1576 TCTC-----TTGCAAGCTCTGCTACTTACAGAAAGTCTGCTACAGACAGAAAGA 1627  
DB 6005 TCTGCCCCAAGACAGACCTGAATCTAGCTGCTAGAGCTGACT CAACGTGAATCA 6063  
QY 1628 TGGCATTTACGGTACTTGAAGAAATAGAGAAATAGAACTATGAAAAATTCATGGAAG 1687  
DB 6064 TGGCGTTTGACAGCACTTGAAGAGTAGACCGAGTGAAGAACTATGACAAAGTTCAATGGAA 6123  
QY 1688 CAATGGCTAAGCCTTACTTTTGAATGCT 1718  
DB 6124 AATGGTAAAGACTTTATTTCTTTTGCT 6154

RESULT 6  
US-11-136-527-3784  
Sequence 3784, Application US/11136527  
Publication No. US20050287570A1  
GENERAL INFORMATION:  
APPLICANT: Wyeth  
APPLICANT: Mounts, William M  
TITLE OF INVENTION: Probe Arrays for Expression Profiling of Rat Genes  
FILE REFERENCE: 031896-041000 (AM101086)  
CURRENT APPLICATION NUMBER: US/11/136,527  
CURRENT FILING DATE: 2005-05-25  
PRIOR APPLICATION NUMBER: US 60/574,294  
PRIOR FILING DATE: 2005-05-26  
NUMBER OF SEQ ID NOS: 362830  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 3784  
LENGTH: 564  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
US-11-136-527-3784

Query Match 2.1%; Score 50; DB 16; Length 564;  
Best Local Similarity 75.6%; Pred. No. 0.15;  
Matches 62; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1615 ACAGACAGAAAGTGCATTTAAAGCTACTTGGAAAAATGAGAAAAATGAACTATGAA 1674  
DB 1 ACAGCTGACATCATGCGCATTTGACACTTGGAAAGTAGACCGAATGAGAACTATGAA 60  
QY 1675 AATTCATGAGCAATGGGTA 1696  
DB 61 AAGTTCATGAGAAATGGGCA 82

RESULT 7  
US-11-136-527-7880  
Sequence 7880, Application US/11136527  
Publication No. US20050287570A1  
GENERAL INFORMATION:  
APPLICANT: Wyeth  
APPLICANT: Mounts, William M

TITLE OF INVENTION: Probe Arrays for Expression Profiling of Rat Genes  
FILE REFERENCE: 031896-041000 (AM101086)  
CURRENT APPLICATION NUMBER: US/11/136,527  
CURRENT FILING DATE: 2005-05-25  
PRIOR APPLICATION NUMBER: US 60/574,294  
PRIOR FILING DATE: 2005-05-26  
NUMBER OF SEQ ID NOS: 362830  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 7880  
LENGTH: 564  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
US-11-136-527-7880

Query Match 2.1%; Score 50; DB 16; Length 564;  
Best Local Similarity 75.6%; Pred. No. 0.15;  
Matches 62; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1615 ACAGACAGAAAGTGCATTTAAAGCTACTTGGAAAAATGAGAAAAATGAGAACTATGAA 1674  
DB 1 ACAGCTGACATCATGCGCATTTGACACTTGGAAAGTAGACCGAATGAGAACTATGAA 60  
QY 1675 AATTCATGAGCAATGGGTA 1696  
DB 61 AAGTTCATGAGAAATGGGCA 82

RESULT 8  
US-10-027-632-5694/C  
Sequence 5694, Application US/10027632  
Publication No. US20020198371A1  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
POLYMORPHISMS IN THE HUMAN GENOME  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5694  
LENGTH: 630  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-5694

Query Match 2.1%; Score 49.8; DB 6; Length 630;  
Best Local Similarity 59.2%; Pred. No. 0.18;  
Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;

QY 1516 TGACCTTAGCCACCATCATCATTAATTTGCTTCCGTATGAGAGCTGTCTATTAAT 1575  
DB 531 TGAACCTTTAGCTTCCACATCACTATGAAAGTTGTTCAAGATTAAGAAATATTAAT 472  
QY 1576 TCTC-----TTGCAAGCTCTGCTACTTACAGAAAGTCTGCTACAGACAGAAAGA 1627  
DB 471 TCTGCCCCAAGACAGACCTGAATCTAGCTGCTAGAGCTGACT CAACGTGAATCA 413  
QY 1628 TGGCATTTACGGTACTTGAAGAAATAGAGAAATAGAACTATGAAAAATTCATGGAAG 1687



```

Db      412 TGGCGTTGACAGCAGCTTGGAGGTAGACCGGAGTGAAGAACTATGACAGTTTCATGGAAA 353
Qy      1688 CAATGGGTAAAGCCTTACTTTTGTGATGCT 1718
Db      352 AAATGGGTAAAGACTTATTTCTTGTGGCT 322

```

RESULT 9

```

US-10-027-632-5694/C
; Sequence 5694, Application US/10027632
; Publication No. US2003020407549
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5694
; LENGTH: 630
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-5694

```

```

Query Match      2.1%; Score 49.8; DB 7; Length 630;
Best Local Similarity 59.2%; Pred. No. 0.18;
Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;

```

```

Qy      1516 TGAGCTTTAGCCAGCATCATCATGTAATGCTTTCCTGATAGCTGTTCTTAAT 1575
Db      531 TGAACCTTAAAGCTTCCACATCAAGTATGAAGTTGTTCAAGATTAAGAAATATTAAT 472
Qy      1576 TCTC-----TTGCAAGCTGCTGCTACTTACCAAGTCTGCTTACAGAGAAAGA 1627
Db      471 TCTGCCCCAAGACAGACCTGAATCTTAAGTCTGCTTACAGAGGCTGACT-CAACTGAAATCA 413
Qy      1628 TGGCATTTAAAGCGTACTTGGAAAAATAGAAAAATGAGAACTATGAAAAATTCATGAG 1687
Db      412 TGGCGTTTGAAGCAGCTTGGAAAGTAGACCGGAGTGAAGAACTATGACAAAGTTTCATGAAA 353
Qy      1688 CAATGGGTAAAGCCTTACTTTTGTGATGCTT 1718
Db      352 AAATGGGTAAAGACTTATTTCTTGTGGCT 322

```

RESULT 10

```

US-10-116-275-303
; Sequence 303, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elan Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgsine, Lisa

```

```

; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; TITLE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell Receptors
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 303
; LENGTH: 5204
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-116-275-303

```

```

Query Match      2.1%; Score 49.8; DB 7; Length 5204;
Best Local Similarity 59.2%; Pred. No. 0.57;
Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;

```

```

Qy      1516 TGAGCTTTAGCCAGCATCATCATGTAATGCTTTCCTGATAGCTGTTCTTAAT 1575
Db      971 TGAACCTTAAAGCTTCCACATCAAGTATGAAGTTGTTCAAGATTAAGAAATATTAAT 1030
Qy      1576 TCTC-----TTGCAAGCTGCTGCTACTTACCAAGTCTGCTTACAGAGAAAGA 1627
Db      1031 TCTGCCCCAAGACAGACCTGAATCTTAAGTCTGCTTACAGAGGCTGACT-CAACTGAAATCA 1089
Qy      1628 TGGCATTTAAAGCGTACTTGGAAAAATAGAAAAATGAGAACTATGAAAAATTCATGAG 1687
Db      1090 TGGCGTTTGAAGCAGCTTGGAAAGTAGACCGGAGTGAAGAACTATGACAAAGTTTCATGAAA 1149
Qy      1688 CAATGGGTAAAGCCTTACTTTTGTGATGCTT 1718
Db      1150 AAATGGGTAAAGACTTATTTCTTGTGGCT 1180

```

RESULT 11

```

US-11-085-792-3
; Sequence 3, Application US/11085792
; Publication No. US20050244864A1
; GENERAL INFORMATION:
; APPLICANT: Kleinfeld, Alan
; TITLE OF INVENTION: FLUORESCENT PROBES FOR HYDROPHOBIC
; TITLE OF INVENTION: ANALYTES
; FILE REFERENCE: PFASC.065PR
; CURRENT APPLICATION NUMBER: US/11/085,792
; CURRENT FILING DATE: 2005-03-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 426
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (1)...(426)
; OTHER INFORMATION: rat intestinal fatty acid binding protein DNA
; OTHER INFORMATION: sequence coding for substitution of alanine for
; OTHER INFORMATION: leucine at position 72
; OTHER INFORMATION: 3' terminus codes for a his tag
US-11-085-792-3

```

```

Query Match      2.1%; Score 49.2; DB 13; Length 426;
Best Local Similarity 81.4%; Pred. No. 0.2;
Matches 57; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```

```

Qy      1627 ATGGCATTTAAAGCTACTTGGAAAAATAGAAAAATGAACTATGAAAAATTCATGAA 1686
Db      1 ATGGCATTTGATGACACTTGGAAAAAGTAGACCGGAATGAACTATGAAAAAGTTTCATGAG 60
Qy      1687 GCATGGGTA 1696
Db      61 AAAATGGGCA 70

```

```
RESULT 12
US-09-925-065A-51803/c
; Sequence 51803, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51803
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-51803

Query Match      2.0%; Score 48; DB 4; Length 546;
Best Local Similarity 68.8%; Pred. No. 0.47;
Matches 66; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY      1623 AAAGTGCATTTAACGGTACTTGAGAAATAGAGAAATGAGAACTATGAAATTCAT 1682
DB      415 AATCATGGCGTTTGACACACTTGAGAGTAGACCGGAGTAAACTATGCAAGTTCAT 356
QY      1683 GGAAGCAATGGGTAAAGCCTTACTTTTGAATGCT 1718
DB      355 GGAATAATGGGTAAAGACTTATTTCTTGTGGCT 320

RESULT 13
US-09-925-065A-51803/c
; Sequence 51803, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51803
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-51803

Query Match      2.0%; Score 48; DB 5; Length 546;
Best Local Similarity 68.8%; Pred. No. 0.47;
Matches 66; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY      1623 AAAGTGCATTTAACGGTACTTGAGAAATAGAGAAATGAGAACTATGAAATTCAT 1682
DB      415 AATCATGGCGTTTGACACACTTGAGAGTAGACCGGAGTAAACTATGCAAGTTCAT 356
QY      1683 GGAAGCAATGGGTAAAGCCTTACTTTTGAATGCT 1718
DB      355 GGAATAATGGGTAAAGACTTATTTCTTGTGGCT 320

RESULT 14
US-10-301-480-153041/c
; Sequence 153041, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 153041
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-153041

Query Match      2.0%; Score 48; DB 12; Length 546;
Best Local Similarity 68.8%; Pred. No. 0.47;
Matches 66; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY      1623 AAAGTGCATTTAACGGTACTTGAGAAATAGAGAAATGAGAACTATGAAATTCAT 1682
DB      415 AATCATGGCGTTTGACACACTTGAGAGTAGACCGGAGTAAACTATGCAAGTTCAT 356
QY      1683 GGAAGCAATGGGTAAAGCCTTACTTTTGAATGCT 1718
DB      355 GGAATAATGGGTAAAGACTTATTTCTTGTGGCT 320

RESULT 15
US-10-301-480-766450/c
; Sequence 766450, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 766450
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-766450

Query Match      2.0%; Score 48; DB 12; Length 546;
Best Local Similarity 68.8%; Pred. No. 0.47;
Matches 66; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
```

Oy 1623 AAAGATGGCATTTAACGGTACTTGAAAAATAGAAAAATGAACTATGAAAAATTCAT 1682  
Db 415 AATCATGGCGTTTGACAGCACTTGAGAGGTAGACCGAGTGAACCTATGACAGTTTCAT 356  
Oy 1683 GGAGCAATGGGTAAAGCCTTACTTTTGAATGCT 1718  
Db 355 GAAAAAATGGGTAAAGACTTATTTCTTGTGGCT 320

Search completed: May 27, 2006, 20:03:04  
Job time : 3254.58 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: May 27, 2006, 18:57:23 ; Search time 42.7971 Seconds  
(without alignments)  
6280.280 Million cell updates/sec

Title: US-10-099-663-1  
Perfect score: 2381  
Sequence: 1 agctcctgcgcagcaaaag5.....atctgaagctcattccag 2381

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 242596 seqs, 56442199 residues

Total number of hits satisfying chosen parameters: 485192

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA\_New:\*  
1: /EMC\_Celerra\_SIDS3/ptocdata/1/pubpna/US09\_NEW\_PUB.seq:\*  
2: /EMC\_Celerra\_SIDS3/ptocdata/1/pubpna/US06\_NEW\_PUB.seq:\*  
3: /EMC\_Celerra\_SIDS3/ptocdata/1/pubpna/US07\_NEW\_PUB.seq:\*  
4: /EMC\_Celerra\_SIDS3/ptocdata/1/pubpna/US08\_NEW\_PUB.seq:\*  
5: /EMC\_Celerra\_SIDS3/ptocdata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
6: /EMC\_Celerra\_SIDS3/ptocdata/1/pubpna/US10\_NEW\_PUB.seq:\*  
7: /EMC\_Celerra\_SIDS3/ptocdata/1/pubpna/US11\_NEW\_PUB.seq:\*  
8: /EMC\_Celerra\_SIDS3/ptocdata/1/pubpna/US60\_NEW\_PUB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49.8	2.1	5204	US-10-524-021-12	Sequence 12, App1
2	38	1.6	1907	US-10-953-349-33745	Sequence 33745, A
3	38	1.6	2244	US-11-217-529-79170	Sequence 79170, A
4	37	1.6	129	US-10-488-619-2835	Sequence 2835, Ap
5	36.6	1.5	394191	US-10-506-549-3	Sequence 3, App1
6	35.8	1.5	4209	US-11-217-529-4068	Sequence 4068, Ap
7	35.2	1.5	2286	US-11-217-529-2925	Sequence 2925, Ap
8	35.2	1.5	3323	US-10-473-173-40	Sequence 40, App1
9	35	1.5	1275	US-11-217-529-2163	Sequence 2163, Ap
10	35	1.5	2439	US-10-953-349-37222	Sequence 37222, A
11	34.6	1.5	111	US-10-488-619-1326	Sequence 1326, Ap
12	34.6	1.5	2565	US-11-315-766-19	Sequence 19, App1
13	34.6	1.5	4692	US-10-511-937-646	Sequence 646, App
14	34.4	1.4	624	US-10-953-349-6180	Sequence 6180, Ap
15	34	1.4	989	US-10-953-349-15742	Sequence 15742, A
16	34	1.4	3426	US-10-511-937-438	Sequence 438, App
17	34	1.4	3877	US-10-196-749-263	Sequence 263, App
18	34	1.4	3877	US-11-101-316-71	Sequence 71, App1
19	33.8	1.4	346	US-10-511-937-529	Sequence 529, App
20	33.8	1.4	560	US-10-953-349-37522	Sequence 37522, A
21	33.8	1.4	750	US-10-953-349-1467	Sequence 1467, Ap
22	33.8	1.4	1146	US-10-953-349-20608	Sequence 20608, A
23	33.8	1.4	1535	US-10-953-349-15855	Sequence 15855, A
24	33.8	1.4	1582	US-11-222-810-10	Sequence 10, App1
25	33.8	1.4	1582	US-11-222-810-12	Sequence 12, App1

C	26	33.8	1.4	1921	6	US-10-953-349-9882	Sequence 9882, Ap
C	27	33.8	1.4	1953	7	US-11-293-697-2048	Sequence 2048, Ap
C	28	33.8	1.4	2205	6	US-10-953-349-9921	Sequence 9921, Ap
C	29	33.8	1.4	3038	7	US-11-293-697-471	Sequence 471, App
C	30	33.8	1.4	4640	6	US-10-196-749-75	Sequence 75, App1
C	31	33.6	1.4	1638	7	US-11-217-529-79683	Sequence 79683, Ap
C	32	33.6	1.4	2320	6	US-10-953-349-8153	Sequence 8153, Ap
C	33	33.6	1.4	2424	7	US-11-217-529-853	Sequence 853, App
C	34	33.6	1.4	4017	6	US-11-326-265-10	Sequence 10, App1
C	35	33.6	1.4	485	7	US-10-196-749-185	Sequence 185, App
C	36	33.4	1.4	485	7	US-11-101-316-43	Sequence 43, App1
C	37	33.4	1.4	883	6	US-10-953-349-15521	Sequence 15521, A
C	38	33.4	1.4	1969	6	US-10-196-749-193	Sequence 193, App
C	39	33.4	1.4	4670	7	US-11-145-307A-29	Sequence 29, App1
C	40	33.2	1.4	644	6	US-10-196-749-213	Sequence 213, App
C	41	33.2	1.4	644	7	US-11-101-316-55	Sequence 55, App1
C	42	33.2	1.4	780	6	US-10-953-349-36666	Sequence 36666, A
C	43	33.2	1.4	1035	7	US-11-217-529-82312	Sequence 82312, A
C	44	33.2	1.4	1566	7	US-11-217-529-79711	Sequence 79711, A
C	45	33.2	1.4	2082	7	US-11-293-697-698	Sequence 698, App

#### ALIGNMENTS

```
RESULT 1
US-10-524-021-12
; Sequence 12, Application US/10524021
; Publication No. US20060099590A1
; GENERAL INFORMATION:
; APPLICANT: NAGOYA INDUSTRIAL SCIENCE RESEARCH INSTITUTE
; APPLICANT: Gifu INTERNATIONAL INSTITUTE OF BIOTECHNOLOGY
; APPLICANT: YAMADA, Yoshiji
; APPLICANT: YOKOTA, Mitsuhiko
; TITLE OF INVENTION: Method for diagnosing a risk of restenosis after percutaneous cor
; FILE REFERENCE: C0200501
; CURRENT APPLICATION NUMBER: US/10/524, 021
; PRIOR FILING DATE: 2005-02-09
; PRIOR APPLICATION NUMBER: JP P2002-233041
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 5204
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-524-021-12

Query Match      2.1% Score 49.8; DB 6; Length 5204;
Best Local Similarity 59.2%; Pred. No. 0.0038;
Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;

QY      1516 TGAAGCTTAAAGCAGCAGCATATCATGTAATTCCTTCTGATAGCCTGTTATTAAT 1575
      |||||
      971 TGAAGCTTAAAGCTTCAATCAGATGTAAGTTGTTCAAGATAAATATATATAAT 1030
      |||||
QY      1576 TCTC-----TTGCAAAAGCTTGTCTACTTACCAAGTTCCTTACAGACGAAGA 1627
      |||||
DB      1031 TCTGCCCCAAGCAGCAGCATGTAATCTAGTGTCTGAGGCTACT-CAACTGAATCA 1089
      |||||
QY      1528 TGGCATTTAAGCGACTTGGAAAAATAGAAAAATGAACTATGAAAAATTCATGGAG 1687
      |||||
DB      1090 TGGCGTTTACAGACACTTGAAGGTAGCCGAGTGAACCTATGACAACTTATGAAA 1149
      |||||
QY      1688 CAATGGGTAAAGCCTTACTTTTGGATGCT 1718
      |||||
DB      1150 AAATGGTAAAGACTTATTTCTTTGGCT 1180
      |||||

RESULT 2
US-10-953-349-33745
; Sequence 33745, Application US/10953349
```

```
Publication No. US20060107345A1
GENERAL INFORMATION:
APPLICANT: ALEXANDROV, NICKOLAI et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
FILE OF INVENTION: ENCLOSED THEREBY
FILE REFERENCE: 2750-1579PUS2
CURRENT APPLICATION NUMBER: US/10/953,349
CURRENT FILING DATE: 2004-09-30
NUMBER OF SEQ ID NOS: 40252
SOFTWARE: PatentIn version 3.3
SEQ ID NO 33745
LENGTH: 1907
TYPE: DNA
ORGANISM: Zea mays subsp. mays
US-10-953-349-33745
```

Query Match 1.6%; Score 38; DB 6; Length 1907;  
Best Local Similarity 62.8%; Pred. No. 2;

Matches 59; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

```
QY 2227 AACTACGAAAAAGTCACATAAATAATGCAACATGATGTCATTGTTGTTCTGCGC 2286
    |||||
DB 1108 AACTATGCAAAAAAAGAAAAAAGCGACGAGAAAGCTTTTCTTTTCTTCTT 1167
    |||||
QY 2287 TTGATGTTAATGATTATTTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2320
    |||||
DB 1168 TTTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1201
    |||||
```

## RESULT 3

US-11-217-529-79170

Sequence 79170, Application US/11217529

Publication No. US2006009612A1

GENERAL INFORMATION:

APPLICANT: SUNTORY LIMITED

APPLICANT: NAKAO, YOSHIHIRO

APPLICANT: NAKAMURA, NORIHISA

APPLICANT: KODAMA, YUKIKO

APPLICANT: FUJIMURA, TOMOKO

APPLICANT: ASHIKARI, TOSHIHIKO

TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS

FILE REFERENCE: S-38-285

CURRENT APPLICATION NUMBER: US/11/217,529

CURRENT FILING DATE: 2005-09-02

PRIOR APPLICATION NUMBER: US 10/932,182

PRIOR FILING DATE: 2004-09-02

NUMBER OF SEQ ID NOS: 197023

SOFTWARE: PatentIn version 3.3

SEQ ID NO 79170

LENGTH: 2244

TYPE: DNA

ORGANISM: Saccharomyces pastorianus

US-11-217-529-79170

Query Match 1.6%; Score 38; DB 7; Length 2244;  
Best Local Similarity 50.5%; Pred. No. 2.1;

Matches 92; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

```
QY 523 ATTATAGTTACTTCTGATTAACCTTTGCTGACGCTTGGAAGAAAAAGAGTTAT 582
    |||||
DB 2002 ATGTATCAAAAGCTGTGGAAGAAAGATCGTGAATAATGCAAAAGAAAGATGAT 2061
    |||||
QY 583 GATTAACCAAGACTGATTAATTCATTAGTTAGCAACCCGTTATCTGTGTGATAG 642
    |||||
DB 2062 AATATACAGCTCTAATCAAGTACGATGATGCTGAAGAACTTGCAATAGTGTCAAAAG 2121
    |||||
QY 643 CAATTCATTTCAGCACTTCAAGATTTTACATTTTGAAGTAAATAGACAGAGATTTGGT 702
    |||||
DB 2122 AAAAAAGAAAGCTGCTATAGCAAGATTTGATCTCTATATAGCAAGCTTTGTT 2181
    |||||
QY 703 GC 704
    |||||
DB 2182 TC 2183
    |||||
```

RESULT 4  
US-10-488-619-2835/C  
Sequence 2835, Application US/10488619  
Publication No. US20060099578A1  
GENERAL INFORMATION:

APPLICANT: Greenlee, Winner and Sullivan, P. C.

TITLE OF INVENTION: Human Mitochondrial DNA Polymorphisms, Haplogroups, Associations &

FILE REFERENCE: 98-01 WO

CURRENT APPLICATION NUMBER: US/10/488,619

CURRENT FILING DATE: 2004-03-01

NUMBER OF SEQ ID NOS: 3040

SOFTWARE: PatentIn version 3.1

SEQ ID NO 2835

LENGTH: 129

TYPE: DNA

ORGANISM: Mus musculus

US-10-488-619-2835

Query Match 1.6%; Score 37; DB 6; Length 129;  
Best Local Similarity 58.7%; Pred. No. 1.2;

Matches 64; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

```
QY 2264 TGTTCATTTTGTCTTCTGCTGATGATTAATTAATTAATTAATTAATTAATTAATTAAT 2323
    |||||
DB 118 TTTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 59
    |||||
QY 2324 CGTGAATGATGAAGAAAGATTAGAGCCGACGATATCGAAGCC 2372
    |||||
DB 58 TTTTATTTAGATTAAGTGTGTTAGGTAGACGAGAAATTTGAATTC 10
    |||||
```

## RESULT 5

US-10-506-549-3

Sequence 3, Application US/10506549

Publication No. US20060100417A1

GENERAL INFORMATION:

APPLICANT: APPLERA CORPORATION

TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,

TITLE OF INVENTION: NOCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,

FILE REFERENCE: CL001361-US

CURRENT APPLICATION NUMBER: US/10/506,549

CURRENT FILING DATE: 2004-09-03

PRIOR APPLICATION NUMBER: 60/361,343

PRIOR FILING DATE: 2002-03-05

NUMBER OF SEQ ID NOS: 4

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 3

LENGTH: 394191

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE: NAME/KEY: misc\_feature

LOCATION: (1)...(394191)

OTHER INFORMATION: n = A,T,C or G

US-10-506-549-3

Query Match 1.5%; Score 36.6; DB 6; Length 394191;  
Best Local Similarity 44.1%; Pred. No. 33;

Matches 153; Conservative 0; Mismatches 194; Indels 0; Gaps 0;

```
QY 1852 TAGAGGGGAATPACAGGAAGAAACATTCGATTTTATTTGATTCGATATCTTATG 1911
    |||||
DB 72108 TAGGGGTGACACAGCGCTGAGCCGCTTTGTATTTTATGTTAGTTAATCTTAAG 72167
    |||||
QY 1912 CATTTAGCTAATTCAGTGAAGGCAATTCAGAGAAATTTAAATAGAAATTAATGTAAG 1971
    |||||
DB 72168 ATATATGATGTTTATTTCTGCTTTTTCATTTAAATATATGTAAGAAATTTTACAT 72227
    |||||
QY 1972 AATATTTATTTTGAATTAACGCTGTTGAAAAATTTACACAGAGGAATTTGCTGCTCCAG 2031
    |||||
```

```
Db 72228 TAGAATTTGTGGCAAACTTTTAAAGACACACCAAGCTTTTATGATAACT 72287
Qy 2032 TTTTGACAGACACATGATTTGAGTCATTTTAACAGCTAGTCTTCTTAAGCTGT 2091
Db 72288 GTAAATGGTGTGATCTCTTTAGTCTCAAAATTTTAGGCTTTTGCTATTTTCCATACC 72347
Qy 2092 ACAACTGCTGTATATGATGATGTAACATTAATCTTCAAGTGGATGATGATTTGTATT 2151
Db 72348 ATTGTATAAAGATATTGTGTGTAATCTTTGCAATTATGATTTTGAATAATAAGTA 72407
Qy 2152 ACAGGCTGACACCTGCTCAGTGAAGGTGAGAAAGTAGAAGTCTT 2198
Db 72408 TAAGACTAAATGATTAATTTTAAAGCTGTGACAAATATCAAGCT 72454
```

RESULT 6  
US-11-217-529-4068/c  
; Sequence 4068, Application US/11217529  
; Publication No. US2006009612A1  
; GENERAL INFORMATION:

APPLICANT: SUNTORY LIMITED  
APPLICANT: NAKAO, YOSHIHIRO  
APPLICANT: NAKAMURA, NORIHISA  
APPLICANT: KODAMA, YUKIKO  
APPLICANT: FUJIMURA, TOMOKO  
APPLICANT: ASHIKARI, TOSHIHIKO  
TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS  
FILE REFERENCE: S-38-285  
CURRENT APPLICATION NUMBER: US/11/217,529  
CURRENT FILING DATE: 2005-09-02  
PRIOR APPLICATION NUMBER: US 10/932,182  
PRIOR FILING DATE: 2004-09-02  
NUMBER OF SEQ ID NOS: 197023  
SOFTWARE: PatentIn version 3.3  
SEQ ID NO 4068  
LENGTH: 4209  
TYPE: DNA  
ORGANISM: Saccharomyces pastorianus  
US-11-217-529-4068

Query Match 1.5%; Score 35.8; DB 7; Length 4209;  
Best Local Similarity 49.2%; Pred. No. 9.4;  
Matches 94; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

```
Qy 1877 CATTCGATTTTATTTGATTCGATATCTTATGCAATTTAGCTAATTCAGTAGAGCA 1936
Db 1867 CATTCCTCTTATTTTACTTGTGATATTTGCTGAGTTCCCTCTCACTGTAAACCCC 1808
Qy 1937 TTCAGCAGAAATTTAATAGATTAATGTAAGAAATTAATTTGATTAAGACTGTTG 1996
Db 1807 TTTCACTTTTCTTCAGTTATTAATAATCTTACATGATGACATGGGTGAGGCCAA 1748
Qy 1997 AAAAATTAACAGAGGAAATGCTGCTCCAGTTTGGACAGACACATGATTTGAG 2056
Db 1747 CATACTCACATACGGAGCTAATTCACATTCCAAATTTCCACAGCCGTTTGAATTTAT 1688
Qy 2057 TCATTTTAACA 2067
Db 1687 GTATTTCTCTCA 1677
```

RESULT 7  
US-11-217-529-2925  
; Sequence 2925, Application US/11217529  
; Publication No. US2006009612A1  
; GENERAL INFORMATION:  
APPLICANT: SUNTORY LIMITED  
APPLICANT: NAKAO, YOSHIHIRO  
APPLICANT: NAKAMURA, NORIHISA  
APPLICANT: KODAMA, YUKIKO  
APPLICANT: FUJIMURA, TOMOKO  
APPLICANT: ASHIKARI, TOSHIHIKO

FILE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS  
FILE REFERENCE: S-38-285  
CURRENT APPLICATION NUMBER: US/11/217,529  
CURRENT FILING DATE: 2005-09-02  
PRIOR APPLICATION NUMBER: US 10/932,182  
PRIOR FILING DATE: 2004-09-02  
NUMBER OF SEQ ID NOS: 197023  
SOFTWARE: PatentIn version 3.3  
SEQ ID NO 2925  
LENGTH: 2286  
TYPE: DNA  
ORGANISM: Saccharomyces pastorianus  
US-11-217-529-2925

Query Match 1.5%; Score 35.2; DB 7; Length 2286;  
Best Local Similarity 53.7%; Pred. No. 10;  
Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

```
Qy 1839 TGAAGAAGCTGTATAGAGGGAATACAGAGAAAGAAACATTTGATTTTATTTGCAATG 1898
Db 1349 TGAGATGCTCGATATAGTATGATACAGAAATACCAAGTATATTTGTTCTGATCGAGAG 1408
Qy 1899 CGATTAATTTTATGCTTATGCTAATTCAGTAGAGGCAATTCAGACAGAAATTAATAGA 1958
Db 1409 ATTAACACCTTACACATCTGGAATTTGCCGAGAGACTTCCACTACAGAACTCAAGACG 1468
Qy 1959 ATTATATGTAAGAAAT 1974
Db 1469 ATTAATCCGTGTATTT 1484
```

RESULT 8  
US-10-473-173-40/c  
; Sequence 40, Application US/10473173  
; Publication No. US20060088823A1  
; GENERAL INFORMATION:

APPLICANT: VAN ANDEL INSTITUTE  
TITLE OF INVENTION: Microarray Gene Expression Profiling in Clear Cell Renal Cell  
FILE REFERENCE: 38345-170094  
CURRENT APPLICATION NUMBER: US/10/473,173  
CURRENT FILING DATE: 2003-09-29  
PRIOR APPLICATION NUMBER: US 60/279,411  
PRIOR FILING DATE: 2001-03-29  
NUMBER OF SEQ ID NOS: 498  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 40  
LENGTH: 3323  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-473-173-40

Query Match 1.5%; Score 35.2; DB 6; Length 3323;  
Best Local Similarity 57.1%; Pred. No. 12;  
Matches 64; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

```
Qy 2217 AAGCTCCTCACTACAGAAAGTCAATTAATAAATGCAAAAGCATGATCTTTTGT 2276
Db 682 AATTTTCATCCACCAAGAAACATCACCAAAATTTCTGGTCAATATGTTTCTTTT 623
Qy 2277 TTTTCTGCTTGATGATTAATGATTAATTTATTTTATTTTATTTTATTTTATTTTATTTT 2328
Db 622 CTTTATCTAGAGATCTTTTATACAGATTAATTAAGGTCCGATGATGCTCTGA 571
```

RESULT 9  
US-11-217-529-2163  
; Sequence 2163, Application US/11217529  
; Publication No. US2006009612A1  
; GENERAL INFORMATION:  
APPLICANT: SUNTORY LIMITED  
APPLICANT: NAKAO, YOSHIHIRO  
APPLICANT: NAKAMURA, NORIHISA

```

; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2163
; LENGTH: 1275
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-2163
```

```

Query Match
Best Local Similarity 1.5%; Score 35; DB 7; Length 1275;
Matches 83; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
```

```

QY 922 AGGAATGACAGAAACAGACATTCCTGAATGGCTATATATAGAGAAATCGTAGAAG 981
DB 794 AACGACAGAGAAAGAAAAGACCAAGAAAGACGAGGATTAAGTAAAGAAAGCTATGATA 853
QY 982 GTGTCTGTAATTTAGACTACTATTAAGAGTGAGACACAGATGAGAAATATCATCGCA 1041
DB 854 AAGCCATCGAAGAAAGAGGAGGAATGTATGGGAAAGAAAGAAATGATGAAACGGC 913
QY 1042 ATTTCGTAGCTCAGACACTAGACTCGAAGCTTTTGAAGTGA 1084
DB 914 AAAAGTCTAGCTCTTCTGAAACCAAGATGATCTGAAAGTGA 956
```

```

RESULT 10
US-10-953-349-37222/c
; Sequence 37222, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 37222
; LENGTH: 2439
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
US-10-953-349-37222
```

```

Query Match
Best Local Similarity 1.5%; Score 35; DB 6; Length 2439;
Matches 62; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
```

```

QY 2264 TGTCTATTTTGTCTTCTCGCTGATGTTAATGATATATATATTTTATTTTATGAG 2323
DB 2375 TTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAA 2316
QY 2324 CGTGAATGTGATGAAAGAAAGTTAGAGCCACGATATCTGAAGC 2370
DB 2315 TATGAATGAGAAATTTGAATGTTGGCTTGCCTGTATGTACAC 2269
```

```

RESULT 11
US-10-488-619-1326
; Sequence 1326, Application US/10488619
; Publication No. US20060099578A1
; GENERAL INFORMATION:
; APPLICANT: Greenlee, Winner and Sullivan, P. C.
; TITLE OF INVENTION: Human Mitochondrial DNA Polymorphisms, Haplogroups, Associations
```

```

; TITLE OF INVENTION: Physiological Conditions, And Genotyping Arrays
; FILE REFERENCE: 98-01 WO
; CURRENT APPLICATION NUMBER: US/10/488,619
; CURRENT FILING DATE: 2004-03-01
; NUMBER OF SEQ ID NOS: 3040
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1326
; LENGTH: 111
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-488-619-1326
```

```

Query Match
Best Local Similarity 1.5%; Score 34.6; DB 6; Length 111;
Matches 58; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
```

```

QY 2261 TGATGTTCTATTTTGTCTTCTCGCTGATGTTAATGATATATATTTTATTTT 2320
DB 5 TGTGGGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 64
QY 2321 AGCGTGAATGTGATGAAAGAAAGTTAGAGCCAC 2357
DB 65 TTTTGTGGGGGTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 101
```

```

RESULT 12
US-11-315-766-19/c
; Sequence 19, Application US/11315766
; Publication No. US20060101544A1
; GENERAL INFORMATION:
; APPLICANT: Butler, Karlene H.
; APPLICANT: Cahoon, Edgar B.
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Ramodu, Omolayo O.
; APPLICANT: Hall, Sarah E.
; TITLE OF INVENTION: Phospholipid:diacylglycerol Acetyltransferases
; FILE REFERENCE: B8146 US NA
; CURRENT APPLICATION NUMBER: US/11/315,766
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US/10/321,802
; PRIOR FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
; LENGTH: 2565
; TYPE: DNA
; ORGANISM: Zea mays
US-11-315-766-19
```

```

Query Match
Best Local Similarity 1.5%; Score 34.6; DB 7; Length 2565;
Matches 52; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
```

```

QY 2264 TGTCTATTTTGTCTTCTCGCTGATGTTAATGATATATATTTTATTTTATGAG 2323
DB 2556 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 2497
QY 2324 CGTGAATGTGATGAAAGAAAGTTAGAGCCAC 2344
DB 2496 TTTTCCGTCTCAAGAGAA 2476
```

```

RESULT 13
US-10-511-937-646/c
; Sequence 646, Application US/10511937
; Publication No. US2006008836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
```



APPLICANT: Morris, MacDonald  
APPLICANT: Rosenberg, Steven  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING  
TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION  
FILE REFERENCE: 506612000104  
CURRENT APPLICATION NUMBER: US/10/511.937  
CURRENT FILING DATE: 2004-10-19  
PRIOR APPLICATION NUMBER: PCT/US2003/012946  
PRIOR FILING DATE: 2003-04-24  
PRIOR APPLICATION NUMBER: US 10/131.831  
PRIOR FILING DATE: 2002-04-24  
PRIOR APPLICATION NUMBER: US 10/325.899  
PRIOR FILING DATE: 2002-12-20  
NUMBER OF SEQ ID NOS: 3117  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 646  
LENGTH: 4692  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-511-937-646

Query Match 1.5%; Score 34.6; DB 6; Length 4692;  
Best Local Similarity 61.8%; Pred. No. 19;  
Matches 55; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 2264 TGTCTATTTTGTCTCTGCTGATGTTAATGATTATTAATTTTATTTTATAG 2323  
DB 4666 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4607  
QY 2324 CGTGAATGTGATGAAGAAGATTAGAG 2352  
DB 4606 TTTGAACAGAAAGAAAGCTTTTATAG 4578

RESULT 14  
US-10-953-349-6180  
Sequence 6180, Application US/10953349  
Publication No. US20060107345A1  
GENERAL INFORMATION:  
APPLICANT: ALEXANDROV, Nikolai et al.  
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
FILE REFERENCE: 2750-1579PUS2  
CURRENT APPLICATION NUMBER: US/10/953.349  
CURRENT FILING DATE: 2004-09-30  
NUMBER OF SEQ ID NOS: 40252  
SOFTWARE: PatentIn version 3.3  
SEQ ID NO 6180  
LENGTH: 624  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-10-953-349-6180

Query Match 1.4%; Score 34.4; DB 6; Length 624;  
Best Local Similarity 48.5%; Pred. No. 9.7;  
Matches 95; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 1816 TGCACATGCGCTGACATTATTTGAAAAGACTTATAGAGGGGAAATCAAGAAAGAAA 1875  
DB 420 TGACATCTCCCTCCCTTTCTTGAAGTATGATTAATTTACTTTCCAAAAA 479  
QY 1876 ACATTGATTTATTTGATTTGATTAATCTTATGCAATTAAGCTTAATTCAGTAGAGC 1935  
DB 480 AGTTGGGCTACACACATCAACCTTACTTTAACTAGCAGCAATTCATCTCAA 539  
QY 1936 ATTCAGCAGAAATTTAATAGAAATATATAGAAATATTTTGAATAGACTGTTT 1995  
DB 540 AAACAAGTAAATCTAACAAGACTAAACACTAATCTTATCTGTTGGCTGCC 599  
QY 1996 GAAAAATTACACAGA 2011  
DB 600 TGAATCGAAAGTAGA 615

RESULT 15  
US-10-953-349-15742/c  
Sequence 15742, Application US/10953349  
Publication No. US20060107345A1  
GENERAL INFORMATION:  
APPLICANT: ALEXANDROV, Nikolai et al.  
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
FILE REFERENCE: 2750-1579PUS2  
CURRENT APPLICATION NUMBER: US/10/953.349  
CURRENT FILING DATE: 2004-09-30  
NUMBER OF SEQ ID NOS: 40252  
SOFTWARE: PatentIn version 3.3  
SEQ ID NO 15742  
LENGTH: 989  
TYPE: DNA  
ORGANISM: Glycine max  
US-10-953-349-15742

Query Match 1.4%; Score 34; DB 6; Length 989;  
Best Local Similarity 61.1%; Pred. No. 15;  
Matches 55; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 2260 ATGATGTTCTATTTTGTCTCTGCTGATGTTAATGATTATTAATTTTATTTT 2319  
DB 843 ATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 784  
QY 2320 TAGCGTGAATGTGATGAAGAAGATTAG 2349  
DB 783 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 754

Search completed: May 27, 2006, 20:04:02  
Job time : 43.7971 secs

THIS PAGE BLANK (USPTO)





repeat_region	5402. .5578	/rpt_family="MIR3"
repeat_region	5380. .5637	/rpt_family="(TA)n"
repeat_region		complement(5638. .5713)
repeat_region		/rpt_family="(LMC4"
	6194. .6321	
	/rpt_family="L2"	
repeat_region	7681. .7843	/rpt_family="MIR"
repeat_region	7849. .7872	
repeat_region	8396. .8447	/rpt_family="CT-rich"
repeat_region		complement(8456. .8777)
repeat_region		/rpt_family="A1u5x"
repeat_region		complement(9376. .9682)
repeat_region	10920. .11043	/rpt_family="A1u4"
repeat_region		/rpt_family="MERSA"
repeat_region	11335. .11363	
repeat_region		/rpt_family="MERSA"
repeat_region		complement(12202. .12268)
repeat_region		/rpt_family="MERSB"
repeat_region	12747. .12861	
repeat_region		/rpt_family="MERS1"
repeat_region	13306. .13396	
repeat_region		/rpt_family="MERSB"
repeat_region	13397. .13422	
repeat_region		/rpt_family="(GGGA)n"
repeat_region	13423. .13488	
repeat_region		/rpt_family="MERSB"
repeat_region	14336. .14471	
repeat_region		/rpt_family="MIR"
repeat_region	15241. .15281	/rpt_family="(TG)n"
repeat_region	16465. .16689	
repeat_region		/rpt_family="(LMC4"
repeat_region	16700. .16766	
repeat_region		/rpt_family="(TA)n"
repeat_region	17390. .18044	
repeat_region		/rpt_family="L2"
repeat_region	18475. .18600	
repeat_region		/rpt_family="L2"
repeat_region	18785. .18937	
repeat_region		/rpt_family="L2"
repeat_region		complement(19502. .19270)
repeat_region		/rpt_family="L2"
repeat_region	19524. .19676	
repeat_region		/rpt_family="(MT1J2"
repeat_region		complement(19677. .19841)
repeat_region		/rpt_family="L2"
repeat_region		complement(20129. .20206)
repeat_region		/rpt_family="MIR"
repeat_region	20779. .21023	
repeat_region		/rpt_family="MIR"
repeat_region		complement(22155. .22575)
repeat_region		/rpt_family="L2"
repeat_region	22873. .22987	
repeat_region		/rpt_family="L2"
repeat_region	23520. .23568	
repeat_region		/rpt_family="MIR"
repeat_region	25325. .25379	
repeat_region		/rpt_family="MIR3"
repeat_region		complement(25438. .25634)
repeat_region		/rpt_family="MIR"
repeat_region	26260. .26378	
repeat_region		/rpt_family="MIR3"
repeat_region	26592. .26622	
repeat_region		/rpt_family="(GC-rich"
repeat_region	28028. .28045	
repeat_region		/rpt_family="(CA)n"
repeat_region	28046. .28077	

	repeat_region	/rpt_family="(CA)n"	28476..28502	
	unsure	/rpt_family="(TTTTA)n"	complement(28639..28644)	
	unsure	/note="<<30 qual SNGL region"	28734..28744	
	unsure	/note="<<30 qual SNGL region"	28768..28774	
	repeat_region	/note="<<30 qual SNGL region"	28773..28793	
	unsure	/rpt_family="GC_rich"	28836..28841	
	unsure	/note="<<30 qual SNGL region"	28855..28867	
	repeat_region	/note="<<30 qual SNGL region"	28974..29317	
	Query Match	9.0%; Score 46; DB 5; Length 16353;		
	Best Local Similarity	45.5%; Pred. No. 0 0043;		
	Matches 163; Conservative 0; Mismatches 195; Indels 0; Gaps 0;			
Oy	23	TATATTGAGTAGAATAAGATCTCACCAAATTAAGAATGGACAAAGCAGAGATTGTG	82	
Dd	48902	TTAAAAGATTGATTAATAATTATTTACTTAATTAATTAATTAATTAACAATTTAAATTTANG	48961	
Oy	83	TTTTATCTGTGGGTAATACGTTTTCACAGTGTAATAAGACCCTCCACACAGTAATA	142	
Dd	48962	GATTACAGATGAATTAATATGCGTATCCAGTGGTAAAAATGGTTAACACATTGAAA	49021	
Oy	143	AGTCCATGCAACAAGAAATGCAATACATCTCTAGTCTCATTTATTTATTTTCATTA	202	
Dd	49022	AATTCATTACTTTTAAAAAGTTAATTTTATTTATTTATTTATTTATTTATTTAGITGAAAAC	49081	
Oy	203	GATAGCCGGTTTTTTACTACACTCAATTAAGATGAACAGATGAATGGATTAGTACTG	262	
Dd	49082	TTTTTATTTTAATTTTCATGGGTACATAGAGCTGCATATATTAATGAGGTACATGATA	49141	
Oy	263	TTTATAAGAAGTAATAAGATPACTATCATCTTTGAGGCATATAGGAGGAGAGAT	322	
Dd	49142	TATTTGATPCAGACATAAAAAAATTCATATTTTAAAAAGATGTTTCCAAGCTATTA	49201	
Oy	323	TCGCAACAGTGCGTTACAATGGGAAAAACAAGTTAACTAAGNAGACCCCCCTCT	380	
Dd	49302	TTTGTATTTAGTATAATAATGAACATGTTTAAAAAGTTAGATCTCACCTCTT	49359	
RESULT 3 BA000016_14				
WCOMMENT	Sequence split into 31 fragments	LOCUS BA000016 Accession BA000016		
	Fragment Name	Begin	End	
	BA000016_00	1	110000	
	BA000016_01	100001	210000	
	BA000016_02	200001	310000	
	BA000016_03	300001	410000	
	BA000016_04	400001	510000	
	BA000016_05	500001	610000	
	BA000016_06	600001	710000	
	BA000016_07	700001	810000	
	BA000016_08	800001	910000	
	BA000016_09	900001	1010000	
	BA000016_10	1000001	1110000	
	BA000016_11	1100001	1210000	
	BA000016_12	1200001	1310000	
	BA000016_13	1300001	1410000	
	BA000016_14	1400001	1510000	
	BA000016_15	1500001	1610000	
	BA000016_16	1600001	1710000	
	BA000016_17	1700001	1810000	
	BA000016_18	1800001	1910000	
	BA000016_19	1900001	2010000	
	BA000016_20	2000001	2110000	
	BA000016_21	2100001	2210000	

BA000016.22 2200001 2310000  
BA000016.23 2300001 2410000  
BA000016.24 2400001 2510000  
BA000016.25 2500001 2610000  
BA000016.26 2600001 2710000  
BA000016.27 2700001 2810000  
BA000016.28 2800001 2910000  
BA000016.29 2900001 3010000  
BA000016.30 3000001 3031430  
Continuation (15 of 31) of BA000016 from base 1400001 (BA000016 Clostridium perfringens)

Query Match 8.9%; Score 45.6; DB 15; Length 110000;  
Best Local Similarity 51.5%; Pred. No. 0.0059;  
Matches 105; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 98 AATACGTTTTCAGTGTATTAAGACCCGCCACAGTAAAGTCCATGCAACA 157  
DB 15707 AAGTTTATTTATTAATTTCTTAATATATTAATCAATCAAGATTAATTCATGTT 15766  
QY 158 AGAAATGCAATACATCTCTAGTCTCATTTATTTTTCATTAGATAGCCGTTT 217  
DB 15767 AATAGATTAATAAGATTAATCTTGACATATATATTTATATATATATTAAT 15826  
QY 218 ACTACACTGAATTAAGATGAAGATGAGTGTAGTGAAGTGTAAAGAGAGT 277  
DB 15827 TCTTATATATGATAGGAGTGTGAAGATGAACATTAATTAATTTTAAAGAAAA 15886  
QY 278 AATTAAGATTAATCATCATTTGA 301  
DB 15887 AATTAATACTAATCTCTCAAGA 15910

RESULT 4  
AR579680 1141 bp DNA linear PAT 14-DEC-2004  
LOCUS AR579680  
DEFINITION Sequence 22 from patent US 6784342.  
ACCESSION AR579680  
VERSION AR579680.1 GI:56583130  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1141)  
AUTHORS Kunze, L. and Clemens, S.  
TITLE Regulation of embryonic transcription in plants  
JOURNAL Patent: US 6784342-A 22 31-AUG-2004;  
The University of British Columbia; Vancouver;  
CAN;

FEATURES  
source Location/Qualifiers  
1..1141  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 8.9%; Score 45.4; DB 2; Length 1141;  
Best Local Similarity 9.9%; Pred. No. 0.01;  
Matches 46; Conservative 197; Mismatches 221; Indels 0; Gaps 0;

QY 12 TCATCATGCTTATTAATGAGTAAATGATTCACCAATTAAGATGCAAGC 71  
DB 110 YBMTNTNKGKTGWRHRYMRABDVTVDHNYTANNAATTCMDKDKRTTMMWKK 169  
QY 72 AGAGATTTGTTTATCTGTGGTAATACGTTTCTCAGTGTATTAAGACCTGC 131  
DB 170 NNATGMDDTKYHMMNNNGCBVTVMVRYKTRDMSBKNNYGMWMMKMSIDVTYYMV 229  
QY 132 CACCAAGATTAAGTCCATGCAACAAGAAATGTCATTAATCTCTTATAGTCTCAT 191  
DB 230 WDMCKRKYRVRWTRGMRNRYVAVBTAHRRYNNGTBAAAYRWYNNNNNNNAKAC 289  
QY 192 TATTTTCATTAGATGAGCGGTTTATTAATGATCAATTAAGATGAAGATGAG 251  
DB 290 KRAKYGMNRABVNSTCTTWSKTTKRTSCVANNCRADANDKHMMKMSAMGVYNN 349

QY 252 GTTAGGACGTGTTTAAGAAGATTAAGATTAATCATCATTTAGGCAATAAG 311  
DB 350 NNNNNNTYKARHBAKMDVWHSAMKMHAAAHYSRKKTBVYRKTVMNNNGTTWKKR 409  
QY 312 GAGGAGAGATTCAGCAACAGTGTCTTCAAGTGAAGAAACAAGTAACTAAGTAC 371  
DB 410 MMATYMKMDMBGTGYNNNNGRTYYGWTNKKMMYYTYSKANNCKMRAMDKTCTHNN 469  
QY 372 CCCCCCTCTTCAACAAGATCAATGCCACAGTTGAGCTTTAGCCGACCATCATGTAA 431  
DB 470 TTTWKKMTYNNCYWKSMTNGKSHRBAAYTYTMMWRRYAHANNNDYWKACCTWYK 529  
QY 432 ATTGCTTCTGATTAAGCTGTTCATTAATTCCTTTGCAAGC 475  
DB 530 YBVCCKMNNYAAAYTSSWNTSRYYRWRKTNNNSMKRSDTRSM 573

RESULT 5  
AX083744 1141 bp DNA linear PAT 28-FEB-2001  
LOCUS AX083744  
DEFINITION Sequence 22 from Patent WO0111061.  
ACCESSION AX083744  
VERSION AX083744.1 GI:13185472  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
synthetic construct  
other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Kunze, L. and Clemens, S.  
TITLE Regulation of embryonic transcription in plants  
JOURNAL Patent: WO 0111061-A 22 15-FEB-2001;  
UNIVERSITY OF BRITISH COLUMBIA (CA)

FEATURES  
source Location/Qualifiers  
1..1141  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
1..1141  
/note="consensus sequence of A.T., L.A., and B.N. FAEI  
promoters"

ORIGIN  
Query Match 8.9%; Score 45.4; DB 2; Length 1141;  
Best Local Similarity 9.9%; Pred. No. 0.01;  
Matches 46; Conservative 197; Mismatches 221; Indels 0; Gaps 0;

QY 12 TCATCATGCTTATTAATGAGTAAATGATTCACCAATTAAGATGCAAGC 71  
DB 110 YBMTNTNKGKTGWRHRYMRABDVTVDHNYTANNAATTCMDKDKRTTMMWKK 169  
QY 72 AGAGATTTGTTTATCTGTGGTAATACGTTTCTCAGTGTATTAAGACCTGC 131  
DB 170 NNATGMDDTKYHMMNNNGCBVTVMVRYKTRDMSBKNNYGMWMMKMSIDVTYYMV 229  
QY 132 CACCAAGATTAAGTCCATGCAACAAGAAATGTCATTAATCTCTTATAGTCTCAT 191  
DB 230 WDMCKRKYRVRWTRGMRNRYVAVBTAHRRYNNGTBAAAYRWYNNNNNNNAKAC 289  
QY 192 TATTTTCATTAGATGAGCGGTTTATTAATGATCAATTAAGATGAAGATGAG 251  
DB 290 KRAKYGMNRABVNSTCTTWSKTTKRTSCVANNCRADANDKHMMKMSAMGVYNN 349  
QY 252 GTTAGGACGTGTTTAAGAAGATTAAGATTAATCATCATTTAGGCAATAAG 311  
DB 350 NNNNNNTYKARHBAKMDVWHSAMKMHAAAHYSRKKTBVYRKTVMNNNGTTWKKR 409  
QY 312 GAGGAGAGATTCAGCAACAGTGTCTTCAAGTGAAGAAACAAGTAACTAAGTAC 371  
DB 410 MMATYMKMDMBGTGYNNNNGRTYYGWTNKKMMYYTYSKANNCKMRAMDKTCTHNN 469  
QY 372 CCCCCCTCTTCAACAAGATCAATGCCACAGTTGAGCTTTAGCCGACCATCATGTAA 431

Db 470 TTWMMKTYNNCKWKSMTNGSKSHBAALVVTYMMWRRYAHANNNDYWKACTWYK 529  
Oy 432 ATTCCTTCCTGATAGCCTGTCATTAATCTCTTGCAGC 475  
Db 530 YBVCCKMNNYAATWTKSSWNTSTRYWTKTNSRMSDTRSM 573

RESULT 6  
AC094376/c  
LOCUS  
DEFINITION Rattus norvegicus clone CH230-3Pl. \*\*\* SEQUENCING IN PROGRESS \*\*\*  
AC094376 238301 bp DNA linear HTG 13-NOV-2002  
AC094376 2 unordered pieces.  
AC094376  
AC094376.11 GI:24942492  
HTG: HTGS PHASE1; HTGS DRAFT; HTGS\_ENRICHED.  
Rattus norvegicus (Norway rat)  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Muridae; Murinae; Rattus.  
1 (bases 1 to 238301)  
Muzny D, Marle, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J.,  
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angiano, D.,  
Anyalibechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,  
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,  
Blawie, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,  
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D Souza, L.,  
Dayla, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,  
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,  
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
Fraser, C. M., Gabriel, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
Georgescu, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,  
Gunnarsson, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,  
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,  
Hollins, B., Howell, S., Hulys, S., Hume, J., Idledid, D., Jackson, A.,  
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
Kapachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
Kowals, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
Lorenz, L., Louised, H., Lozano, R. J., Lu, X., Ma, J.,  
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,  
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,  
Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,  
Milošević, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,  
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
Nwaokeme, O., Okwona, G., Olamunsgoon, A., Pal, S., Parks, K.,  
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,  
Plopper, F., Poldexter, A., Popovic, D., Prims, E., Pu, L.-L.,  
Piaz, M., Quiroz, J., Rachlin, B., Reeves, K., Regier, M. A., Reigh, R.,  
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,  
Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,  
Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,  
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smales, D.,  
Snead, A., Sodergren, E., Song, X.-Z., Sorrell, R., Sosa, J.,  
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabors, P., Taylor, C.,  
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,  
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,  
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,  
Williams, G., Wilson, R., Wleczyk, R., Woodson, H., Wortley, K.,  
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von  
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,  
Weinstock, G. and Gibbs, R. A.  
Direct Submission  
2 (bases 1 to 238301)

AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Wortley, K. C.  
Direct Submission  
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 238301)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (13-NOV-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Nov 13, 2002 this sequence version replaced gi:23265811.  
The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(<http://www.hgsc.bcm.tmc.edu/projects/atlantis/>). Each contig described  
in the feature table below represents a scaffold in the Atlas  
assembly (a 'contig-scaffold'). Within each contig-scaffold,  
individual sequence contigs are ordered and oriented, and separated  
by sized gaps filled with Ns to the estimated size. The sequence  
may extend beyond the ends of the clone and there may be sequence  
contigs within a contig-scaffold that consist entirely of whole  
genome shotgun sequence reads. Both end sequences and whole genome  
shotgun sequence only contigs will be indicated in the feature  
table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: GAGE  
Center clone name: CH230-3Pl  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 229728 bases at least Q40  
Consensus quality: 232177 bases at least Q30  
Consensus quality: 233540 bases at least Q20  
Estimated insert size: 237791; sum-of-contigs estimation  
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 237092: contig of 237092 bp in length  
\* 237093 237192: gap of unknown length  
\* 237193 238301: contig of 1109 bp in length.  
Location/Qualifiers  
1. 238301  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-3Pl"  
237093..237192  
/estimated\_length=unknown

ORIGIN  
Query Match 8.9%; Score 45.4; DB 12; Length 238301;  
Best Local Similarity 53.0%; Pred. No. 0.0064;  
Matches 97; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Oy 173 ATTCCTTCCTGATAGCCTGTCATTAATCTCTTGCAGC 232  
Db 152787 ATTCCTTCCTGATAGCCTGTCATTAATCTCTTGCAGC 232  
Oy 233 AGATGAACAGATGAATGCTAGTCTGTTTAAAGAGATTAAGATCAATC 292

Db 152727 GAGTCAGCATTTAAATGTGATAGTATCTCCCAAAAAATTAATATAAAAT 152668  
QY 293 ATCATTTAGGCATTAAGGAGGAGAGATTCAGCAACAGTGTCTTACAGTGGAAAA 352  
Db 152667 GGAATATGAGCAATCAGAAAGAACACATGAAACCACTTGACATAGAAATCAAAA 152608  
QY 353 CAA 355  
Db 152607 GAA 152605

RESULT 7  
CT030043/c 161172 bp DNA 1linear HTG 04-OCT-2005  
LOCUS Danio rerio chromosome 9 clone CH211-142P19, \*\*\* SEQUENCING IN  
DEFINITION  
ACCESSION CT030043  
VERSION CT030043.2 GI:77019455  
KEYWORDS HTG; HTGS PHASE1.  
SOURCE Danio rerio (zebrafish)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.  
1 (bases 1 to 161172)  
Sims,S.  
Direct Submission  
Submitted (03-OCT-2005) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
fish-helpe@sanger.ac.uk Clone requests:  
http://www.sanger.ac.uk/projects/D\_rerio/fags.shtml#dataight  
On Oct 4, 2005 this sequence version replaced gi:76873691.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: zfish-helpe@sanger.ac.uk  
----- Project Information  
Center project name: ZC142P19  
----- Summary Statistics  
Assembly program: XGAP4; version 4.5  
Chemistry: Dye-terminator; 100% of reads  
Consensus quality: 158423 bases at least Q40  
Consensus quality: 158998 bases at least Q20  
Consensus quality: 159397 bases at least Q20  
Insert size: 160372; sum-of-coverage  
Insert size: 164408; 2.0% error; agarose-fp  
Quality coverage: 6.81x in Q20 bases; sum-of-coverage  
coverage: 6.95x in Q20 bases; agarose-fp  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 9 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1  
6957: contig of 6957 bp in length  
6958 7057: gap of 100 bp  
7058 13672: contig of 6615 bp in length  
13673 13772: gap of 100 bp  
13773 48822: contig of 35050 bp in length  
48823 48922: gap of 100 bp  
48923 107378: contig of 58456 bp in length  
107379 107478: gap of 100 bp  
107479 117380: contig of 9902 bp in length  
117381 117480: gap of 100 bp  
117481 124960: contig of 7480 bp in length  
124961 125060: gap of 100 bp  
125061 133128: contig of 8066 bp in length  
133129 133228: gap of 100 bp  
133229 152597: contig of 19369 bp in length

FEATURES  
\* 152598 152697: gap of 100 bp  
\* 152698 161172: contig of 8475 bp in length.  
source  
1. 161172  
/organism="Danio rerio"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/chromosome="9"  
/clone="CH211-142P19"  
/clone\_11b="CHOR1-211"  
1. 6957  
/note="assembly\_fragment:00258  
fragment\_chain:1  
clone\_end:SP6  
vector\_side:left"  
7058\_13672  
/note="assembly\_fragment:00109  
fragment\_chain:1"  
13773\_48822  
/note="assembly\_fragment:00767  
fragment\_chain:2"  
48923\_107378  
/note="assembly\_fragment:00189  
fragment\_chain:2"  
107479\_117380  
/note="assembly\_fragment:00436  
fragment\_chain:2"  
117481\_124960  
/note="assembly\_fragment:00182  
fragment\_chain:2"  
125061\_133128  
/note="assembly\_fragment:00046"  
133229\_152597  
/note="assembly\_fragment:00533  
fragment\_chain:3"  
152698\_161172  
/note="assembly\_fragment:00339  
fragment\_chain:3  
clone\_end:17  
vector\_side:right"  
ORIGIN  
Query Match 8.8%; Score 44.8; DB 12; Length 161172;  
Best Local Similarity 58.1%; Pred. No. 0.01; 57; Indels 0; Gaps 0;  
Matches 79; Conservative 0; Mismatches 57; Indels 0; Gaps 0;  
QY 232 AAGATGACACAGTAATAGCGTTAGTACGTTTATTAAGAAGATTAAGATCTAT 291  
Db 7209 AAGATGAGTAAAGATGAGGTTTATTTATTAAGAAAGATGAATGAATGATAT 7150  
QY 292 CATCATTTAGGCATTAAGGAGGAGAGATTCAGCAACAGTGTCTTACAGTGGAAA 351  
Db 7149 AATTAAAGGTGAGTAATAGCGTGGGGGAAATTGAAATAATATATGCTGATGATAAAA 7090  
QY 352 ACAAGTTAACTTAAAG 367  
Db 7089 AAAAAATGATGAATAG 7074

RESULT 8  
AC026485/c 69335 bp DNA 1linear HTG 22-MAR-2000  
LOCUS Homo sapiens clone RP11-24F22, LOW-PASS SEQUENCE SAMPLING.  
DEFINITION AC026485  
AC026485.1 GI:7280315  
VERSION  
KEYWORDS HTG; HTGS PHASE0.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 69335)  
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.



TITLE Homo sapiens, clone RP11-24F22  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 69335)  
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Baerlein, V., Bede, P.,  
Boguslavsky, L., Bouckgalter, B., Brown, A., Burkett, G.,  
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,  
Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S.,  
Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
Grand-pierre, N., Grant, G., Hages, B., Heatford, A., Horton, L.,  
Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
Klein, J., Lacroque, K., Lamazares, R., Landers, T., Lehocck, J.,  
Levine, R., Liu, C., Liu, G., Locke, K., MacDonald, P., Marquis, N.,  
McCarthy, M., McGowan, P., McGuirk, A., McKernan, K., McPheters, R.,  
Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mienna, V., Morrow, J.,  
Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N.,  
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Testafay, S., Theodore, J., Titrrell, A., Travers, M., Triggilio, J.,  
Vasilev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker:  
Smit, A.P.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
-----  
Center: Whitehead Institute/ MIT Center for Genome Research  
Genome Center  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
-----  
Project Information  
Center project name: L4446  
Center clone name: 24\_F\_22  
-----  
\* NOTE: This record contains 88 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.  
\* 1  
\* 709: contig of 709 bp in length  
\* 710 809: gap of 100 bp  
\* 810 1511: contig of 702 bp in length  
\* 1512 1611: gap of 100 bp  
\* 2309: contig of 698 bp in length  
\* 2310 2409: gap of 100 bp  
\* 2410 3108: contig of 699 bp in length  
\* 3109 3208: gap of 100 bp  
\* 3209 3893: contig of 685 bp in length  
\* 3894 3993: gap of 100 bp  
\* 3994 4686: contig of 693 bp in length  
\* 4687 4786: gap of 100 bp  
\* 4787 5482: contig of 696 bp in length  
\* 5483 5582: gap of 100 bp  
\* 5583 6292: contig of 710 bp in length  
\* 6293 6392: gap of 100 bp  
\* 6393 7092: contig of 700 bp in length  
\* 7093 7192: gap of 100 bp  
\* 7193 7865: contig of 673 bp in length  
\* 7866 7965: gap of 100 bp  
\* 7966 8657: contig of 692 bp in length  
\* 8657 8757: gap of 100 bp  
\* 8758 9444: contig of 687 bp in length  
\* 9445 9544: gap of 100 bp  
\* 9545 10237: contig of 693 bp in length  
\* 10238 10337: gap of 100 bp  
\* 10338 11014: contig of 677 bp in length  
\* 11015 11114: gap of 100 bp  
\* 11115 11778: contig of 664 bp in length  
\* 11779 11878: gap of 100 bp  
\* 11879 12555: contig of 677 bp in length  
\* 12556 12655: gap of 100 bp  
\* 12656 13360: contig of 705 bp in length  
\* 13361 13460: gap of 100 bp  
\* 13461 14163: contig of 703 bp in length  
\* 14164 14263: gap of 100 bp  
\* 14264 14934: contig of 671 bp in length  
\* 14935 15034: gap of 100 bp  
\* 15035 15720: contig of 686 bp in length  
\* 15721 15820: gap of 100 bp  
\* 15821 16514: contig of 694 bp in length  
\* 16515 16614: gap of 100 bp  
\* 16615 17304: contig of 690 bp in length  
\* 17305 17404: gap of 100 bp  
\* 17405 18093: contig of 689 bp in length  
\* 18094 18193: gap of 100 bp  
\* 18194 18891: contig of 698 bp in length  
\* 18892 18991: gap of 100 bp  
\* 18992 19672: contig of 681 bp in length  
\* 19673 19772: gap of 100 bp  
\* 19773 20455: contig of 683 bp in length  
\* 20456 20555: gap of 100 bp  
\* 20556 21339: contig of 684 bp in length  
\* 21340 22026: contig of 687 bp in length  
\* 22027 22126: gap of 100 bp  
\* 22127 22826: contig of 700 bp in length  
\* 22827 22926: gap of 100 bp  
\* 22927 23625: contig of 699 bp in length  
\* 23626 23725: gap of 100 bp  
\* 23726 24393: contig of 668 bp in length  
\* 24394 24493: gap of 100 bp  
\* 24494 25181: contig of 688 bp in length  
\* 25182 25281: gap of 100 bp  
\* 25282 25984: contig of 703 bp in length  
\* 25985 26084: gap of 100 bp  
\* 26085 26765: contig of 681 bp in length  
\* 26766 27573: gap of 100 bp  
\* 27574 27673: contig of 708 bp in length  
\* 27674 28453: gap of 100 bp  
\* 28454 29142: contig of 680 bp in length  
\* 29143 29242: gap of 100 bp  
\* 29243 29827: contig of 685 bp in length  
\* 29828 30027: gap of 100 bp  
\* 30028 30719: contig of 692 bp in length  
\* 30720 30819: gap of 100 bp  
\* 30820 31514: contig of 695 bp in length  
\* 31515 31614: gap of 100 bp  
\* 31615 32308: contig of 694 bp in length  
\* 32309 32408: gap of 100 bp  
\* 32409 33094: contig of 686 bp in length  
\* 33095 33194: gap of 100 bp  
\* 33195 33895: contig of 701 bp in length  
\* 33896 33995: gap of 100 bp  
\* 33996 34705: contig of 710 bp in length  
\* 34706 34805: gap of 100 bp  
\* 34806 35504: contig of 699 bp in length  
\* 35505 35604: gap of 100 bp  
\* 35605 36291: contig of 687 bp in length  
\* 36292 36391: gap of 100 bp  
\* 36392 37090: contig of 695 bp in length  
\* 37091 37190: gap of 100 bp  
\* 37191 37881: contig of 691 bp in length  
\* 37882 37981: gap of 100 bp

```

*      37982      38671: contig of 690 bp in length
*      38672      38771: gap of 100 bp
*      38772      39465: contig of 694 bp in length
*      39465      39565: gap of 100 bp
*      39565      40255: contig of 690 bp in length
*      40255      40355: gap of 100 bp
*      40355      41052: contig of 697 bp in length
*      41052      41152: gap of 100 bp
*      41152      41847: contig of 695 bp in length
*      41847      41947: gap of 100 bp
*      41947      42630: contig of 683 bp in length
*      42630      42730: gap of 100 bp
*      42730      43417: contig of 687 bp in length
*      43417      43517: gap of 100 bp
*      43517      44185: contig of 668 bp in length
*      44185      44285: gap of 100 bp
*      44285      44986: contig of 701 bp in length
*      44986      45086: gap of 100 bp
*      45086      45766: contig of 680 bp in length
*      45766      45866: gap of 100 bp
*      45866      46530: contig of 664 bp in length
*      46530      46630: gap of 100 bp
*      46630      47323: contig of 693 bp in length
*      47323      47423: gap of 100 bp
*      47423      48022: contig of 599 bp in length
*      48022      48122: gap of 100 bp
*      48122      48778: contig of 656 bp in length
*      48778      48878: gap of 100 bp
*      48878      49579: contig of 701 bp in length
*      49579      49679: gap of 100 bp
*      49679      50387: contig of 708 bp in length
*      50387      50487: gap of 100 bp
*      50487      51192: contig of 705 bp in length
*      51192      51292: gap of 100 bp
*      51292      51980: contig of 688 bp in length
*      51980      52080: gap of 100 bp
*      52080      52784: contig of 704 bp in length
*      52784      52884: gap of 100 bp
*      52884      53572: contig of 688 bp in length
*      53572      53672: gap of 100 bp
*      53672      54365: contig of 693 bp in length
*      54365      54465: gap of 100 bp

```

Query Match Best Local Similarity 8.7%; Score 44.6; DB 12; Length 69335; Matches 145; Conservative 0; Mismatches 149; Indels 1; Gaps 1;

```

Qy      49 CACCAATTTAGATGACAAAGACAGATTTGTTTATCTGTTGGTAAATACGTTT 108
Db      62189 CAGGCACACGCGGACATGTATATGATGACTGACCAATGTGCACATGTA 62130
Qy      109 CTCAGTTGTAAGAGCCCTCCACCACTATTAAGTCTTATGCAACAAAGAAATGTC 168
Db      62129 CCTTAAACCTTAAAGTATTAATTAATTAATTAATTAATTAATTAATTAAT 62070
Qy      169 ATACATTTCTTACTCTCATATATTTTCTATGATAGCGGTTTTTTTCTCAACTCA 228
Db      62069 TAAATTTCTTAATCTT-ACCTTCAGTGCATCTCGAAGCTTTTCTTAAATATATGA 62011
Qy      229 AATTAAGTAAACAGATGATGAGTATGATGATGATGATGATGATGATGATGATG 288
Db      62010 CTTTATTTTCAACAGTGTATGAAATATGAGCTTATTTTAAAGAAACAGTATTTATTTG 61951
Qy      289 TATCATATTTTGAAGCAATAGGAGGAGGAGGAGATTCAGCAACAGTGTCTTACA 343
Db      61950 AAAAAACCTTATGATGTGAAGAGAGCAAGATGTGAACCAAGCCGTGATATAAA 61896

```

RESULT 9 AC127531 52915 bp DNA linear HTG 17-JUL-2002  
 AC127531 Homo sapiens chromosome 11 clone CTD-2362F20 map 11, LOW-PASS  
 DEFINITION SEQUENCE SAMPLING.

ACCESSION AC127531  
 VERSION AC127531.1 GI:21886951  
 KEYWORDS HTG; HTGS\_PHASE0.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo  
 1 (bases 1 to 52915)  
 Birren, B., Nusbaum, C., and Lander, E.  
 Homo sapiens chromosome 11, clone CTD-2362F20  
 Unpublished  
 2 (bases 1 to 52915)

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 Birren, B., Nusbaum, C., Lander, E., Allen, A., Allen, N., Anderson, S., Barina, N., Baerlein, V., Bloom, T., Boguslavsky, L., Bouhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardina, S., Gord, S., Graham, L., Grand-pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrum, J., Menus, L., Milhova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retz, R., Rise, C., Rogov, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Teblaye, S., Theodores, J., Topham, K., Travers, M., Vasiliyev, H., Vei, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zemke, L., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (17-JUL-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 All repeats were identified using RepeatMasker:  
 Smit, A. F. A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

COMMENT  
 TITLE  
 JOURNAL

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: L27588  
 Center clone name: 2362\_F\_20

\* NOTE: This record contains 64 individual  
 \* sequencing reads that have not been assembled into  
 \* contigs. Runs of N are used to separate the reads  
 \* and the order in which they appear is completely  
 \* arbitrary. Low-pass sequence sampling is useful for  
 \* identifying clones that may be gene-rich and allows  
 \* overlap relationships among clones to be deduced.  
 \* However, it should not be assumed that this clone  
 \* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved.  
 \* 1  
 \* 689 688: contig of 688 bp in length  
 \* 789 788: gap of 100 bp  
 \* 1509 1509: contig of 721 bp in length  
 \* 1510 1509: gap of 100 bp  
 \* 1610 2354: contig of 745 bp in length  
 \* 2355 2454: gap of 100 bp  
 \* 2455 3158: contig of 704 bp in length  
 \* 3159 3258: gap of 100 bp  
 \* 3259 3989: contig of 731 bp in length  
 \* 3990 4089: gap of 100 bp  
 \* 4090 4825: contig of 736 bp in length  
 \* 4826 4925: gap of 100 bp  
 \* 4926 5657: contig of 732 bp in length  
 \* 5658 5757: gap of 100 bp  
 \* 5758 6489: contig of 732 bp in length  
 \* 6490 6589: gap of 100 bp

```
* 6590 7298: contig of 709 bp in length
* 7299 7398: gap of 100 bp
* 7399 8132: contig of 734 bp in length
* 8133 8232: gap of 100 bp
* 8233 8962: contig of 730 bp in length
* 8963 9062: gap of 100 bp
* 9063 9790: contig of 728 bp in length
* 9791 9890: gap of 100 bp
* 9891 10625: contig of 735 bp in length
* 10626 10725: gap of 100 bp
* 10726 11456: contig of 731 bp in length
* 11457 11556: gap of 100 bp
* 11557 12283: contig of 727 bp in length
* 12284 12383: gap of 100 bp
* 12384 13130: contig of 747 bp in length
* 13131 13230: gap of 100 bp
* 13231 13983: contig of 753 bp in length
* 13984 14083: gap of 100 bp
* 14084 14832: contig of 749 bp in length
* 14833 14932: gap of 100 bp
* 14933 15661: contig of 729 bp in length
* 15662 15761: gap of 100 bp
* 15762 16482: contig of 721 bp in length
* 16483 16582: gap of 100 bp
* 16583 17298: contig of 716 bp in length
* 17299 17398: gap of 100 bp
* 17399 18144: contig of 746 bp in length
* 18145 18244: gap of 100 bp
* 18245 18979: contig of 735 bp in length
* 18980 19079: gap of 100 bp
* 19080 19797: contig of 718 bp in length
* 19798 19897: gap of 100 bp
* 19898 20621: contig of 724 bp in length
* 20622 20721: gap of 100 bp
* 20722 21470: contig of 749 bp in length
* 21471 21570: gap of 100 bp
* 21571 22316: contig of 746 bp in length
* 22317 22416: gap of 100 bp
* 22417 23125: contig of 709 bp in length
* 23126 23225: gap of 100 bp
* 23226 23968: contig of 743 bp in length
* 23969 24068: gap of 100 bp
* 24069 24794: contig of 726 bp in length
* 24795 24894: gap of 100 bp
* 24895 25621: contig of 727 bp in length
* 25622 25721: gap of 100 bp
* 25722 26453: contig of 732 bp in length
* 26454 26553: gap of 100 bp
* 26554 27300: contig of 747 bp in length
* 27301 27400: gap of 100 bp
* 27401 28125: contig of 725 bp in length
* 28126 28225: gap of 100 bp
* 28226 28974: contig of 749 bp in length
* 28975 29074: gap of 100 bp
* 29075 29816: contig of 742 bp in length
* 29817 29916: gap of 100 bp
* 29917 30663: contig of 747 bp in length
* 30664 30763: gap of 100 bp
* 30764 31497: contig of 734 bp in length
* 31498 31597: gap of 100 bp
* 31598 32331: contig of 734 bp in length
* 32332 32431: gap of 100 bp
* 32432 33159: contig of 728 bp in length
* 33160 33259: gap of 100 bp
* 33260 33988: contig of 729 bp in length
* 33989 34088: gap of 100 bp
* 34089 34835: contig of 747 bp in length
* 34836 34935: gap of 100 bp
* 34936 35671: contig of 736 bp in length
* 35672 35771: gap of 100 bp
* 35772 36505: contig of 734 bp in length
* 36506 36605: gap of 100 bp
* 36606 37286: contig of 681 bp in length
```

```
* 37287 37386: gap of 100 bp
* 37387 38115: contig of 729 bp in length
* 38116 38215: gap of 100 bp
* 38216 38944: contig of 729 bp in length
* 38945 39044: gap of 100 bp
* 39045 39765: contig of 721 bp in length
* 39766 39865: gap of 100 bp
* 39866 40599: contig of 734 bp in length
* 40600 40699: gap of 100 bp
* 40700 41441: contig of 742 bp in length
* 41442 41541: gap of 100 bp
* 41542 42157: contig of 616 bp in length
* 42158 42257: gap of 100 bp
* 42258 42989: contig of 732 bp in length
* 42990 43089: gap of 100 bp
* 43090 43805: contig of 716 bp in length
* 43806 43905: gap of 100 bp
* 43906 44654: contig of 749 bp in length
* 44655 44754: gap of 100 bp
* 44755 45498: contig of 744 bp in length
* 45499 45598: gap of 100 bp
* 45599 46335: contig of 737 bp in length
* 46336 46435: gap of 100 bp
* 46436 47149: contig of 714 bp in length
* 47150 47249: gap of 100 bp
* 47250 47980: contig of 731 bp in length
* 47981 48080: gap of 100 bp
* 48081 48819: contig of 739 bp in length
* 48820 48919: gap of 100 bp
* 48920 49646: contig of 727 bp in length
* 49647 49746: gap of 100 bp
* 49747 50473: contig of 727 bp in length
* 50474 50573: gap of 100 bp
* 50574 51269: gap of 100 bp
* 51269 51368: gap of 100 bp
* 51369 52094: contig of 726 bp in length
* 52095 52194: gap of 100 bp
* 52195 52915: contig of 721 bp in length.

FEATURES
source
1..52915
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="11"
    /map="11"
    /clone="CTD-2362F20"
    /clone_11b="CITD1 Human BAC"
    /estimated_length=100
    /estimated_length=100
    /estimated_length=100

Query Match      8.7%; Score 44.4; DB 12; Length 52915;
Best Local Similarity 50.0%; Pred.No. 0.015;
Matches 141; Conservative 0; Mismatches 136; Indels 5; Gaps 1;

QY      19  ATGCTATTATGTGAGTGAATAGTATCTTCACCAATTAAGATGACAAAGCAAGAT 78
Db      39896  ATTCGCGAGTGAAGATGACAAATCTTAAGACAAAGTAAAGTCCAGAGAGAT 39955
QY      79  TGTGTTATCTGTGGTAAATAGTTCCTCCAGTGTGTAAGACCTCCACCAAGT 138
Db      39956  AAATTCAGTCTATTCACCTAAATATTTCTATCAAGAACTGGAAAACTTCCA----- 40010
QY      139  ATAAAGCTTAATGCAACAAAGAAATGTCAATTAATCTCTTAAGTCAATTAATTTTC 198
Db      40011  AATAAGCTCACTGAACATAAAGAAATTAAGAAATTAATGTTGTAAGAAAGTGAAGTCTT 40070
QY      199  ATTGATATGCCGCTTTTACTACACTCAATTAAGATGAACAGATGAATGGTATG 258
Db      40071  AAAAGATGAGATGAGAACTACCGCTTCAGTCTGAGAAAGTGAAGATGAGACAA 40130
QY      259  ACTGTTATAAGAAAGTAAATTAAGATATCATTCATTTG 300
```

RESULT_10	Db	40131	ACTCTACCAACTACAGAGAGAGTGCATCTTGTGTACCTTGG	40172
LOCUS CR956401				
DEFINITION Danto rerio clone DKEY-5711, WORKING DRAFT SEQUENCE, 11 unordered pieces.				
ACCESSION CR956401		206489 bp	DNA	linear
VERSION CR956401.3				
KEYWORDS HTG; HTGS_PHASE1; HTGS DRAFT; HTGS_FULLTOP.				
SOURCE Danto rerio (zebrafish)				
ORGANISM Danto rerio				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danto.				
1 (bases 1 to 206489)				
McLaren,S.				
Direct Submission				
Submitted (29-JAN-2006) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: http://www.sanger.ac.uk/Projects/D_rerio/fags.shtml#dataight				
On Jan 30, 2006 this sequence version replaced gi:67509286.				
Genome Center				
Center: Wellcome Trust Sanger Institute				
Center code: SC				
Web site: http://www.sanger.ac.uk				
Contact: zfish-help@sanger.ac.uk				
Project Information				
Center project name: ZK5711				
Summary Statistics				
Assembly program: XGAP4; version 4.5				
Chemistry: Dye-terminator; 100% of reads				
Consensus quality: 202739 bases at least Q40				
Consensus quality: 203666 bases at least Q30				
Consensus quality: 204251 bases at least Q20				
Insert size: 205489; sum-of-contrigs				
Insert size: 249713; 4.3% error; agarose-fp				
Quality coverage: 12.96x in Q20 bases; sum-of-contrigs Quality coverage: 10.86x in Q20 bases; agarose-fp				
NOTE: This is a 'working draft' sequence. It currently consists of 11 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.				
1 36677: contig of 36677 bp in length				
36678 36777: gap of 100 bp				
36778 79240: contig of 42463 bp in length				
79241 79340: gap of 100 bp				
79341 86787: contig of 7447 bp in length				
86788 86887: gap of 100 bp				
86888 93958: contig of 7071 bp in length				
93959 94058: gap of 100 bp				
94059 101325: contig of 7267 bp in length				
101326 101425: gap of 100 bp				
101426 106157: contig of 4732 bp in length				
106158 106257: gap of 100 bp				
106258 110360: contig of 4103 bp in length				
110361 110460: gap of 100 bp				
110461 116362: contig of 5902 bp in length				
116363 116462: gap of 100 bp				
116463 176493: contig of 60031 bp in length				
176494 176593: gap of 100 bp				
176594 180881: contig of 4288 bp in length				
180882 180981: gap of 100 bp				
180982 206489: contig of 2508 bp in length.				
Location/Qualifiers				

```

source
1..206489
/organism="Danio rerio"
/md_type="genomic DNA"
/db_xref="taxon:7955"
/clone_id="DKEY-5711"
/clone_id="DanioKey"
1..36677
/note="assembly_fragment:00001
fragment_chain:1"
36778..79240
/note="assembly_fragment:00713
fragment_chain:1"
79341..86787
/note="assembly_fragment:03265
fragment_chain:1"
86888..93958
/note="assembly_fragment:03188
fragment_chain:1"
94059..101325
/note="assembly_fragment:02996
fragment_chain:1"
101426..106157
/note="assembly_fragment:02967
fragment_chain:1"
106258..110360
/note="assembly_fragment:03032
fragment_chain:1"
110461..116362
/note="assembly_fragment:03077
fragment_chain:1"
116463..176493
/note="assembly_fragment:01663"
176594..180881
/note="assembly_fragment:03130.0"
180982..206489
/note="assembly_fragment:03345"

```

ORIGIN	Query Match	8.7% ;	Score 44.4 ;	DB 12 ;	Length 206489 ;
	Best Local Similarity	43.5% ;	Pred. No. 0.013 ;		
	Matches 117 ;	Conservative	0 ;	Mismatches 152 ;	Indels 0 ;
				Gaps	0 ;
QY	55	ATTGAAGTGGACAAAGCAGAGATTGGTGTTCCTATCTGTGGGTAAACGCTTTCCACG	114		
Db	110503	AATTAATAAAAAAAAAAANTTTNTTTNTTTNTTTTAAAAAATAATATTTTNTNN	110562		
QY	115	TTGTATTAAGACCCCTCCACGATTAATAAGTCATGCAACAAAGAAATGTCATATCAT	174		
Db	110563	NNNNNAATAATTTTAAATAAANNNNNBAAATTTTAAATAAANNTAAATNAAAAANNA	110622		
QY	175	TCTCTAGTCTCATTTATTTTCATATGATAGCCGGTTTCTACTCAACTGAAATPAG	234		
Db	110623	ATATTTTTTTTTTATTTATTTTNTTAAANNNNTAATNTTTTAAATNTNTTTTATTTAAA	110682		
QY	235	ATGACAGCAATGAATGGGTAGCTGTTTATAAGAGATATAAAGATACATCAT	294		
Db	110683	AAAAAAAAAATAATTAATAAATTTTTTTTTTAAATATATTTTAAAAAATAATTTTTTTT	110742		
QY	295	CATTGAGCATPAGGAGGAGGAGATT	323		
Db	110743	TTTATATATAAAAAAAAAATTAATANAATATTT	110771		
RESULT 11					
PFMAL13PB/c					
LOCUS	PFMAL13PB	83110 bp	DNA	linear	HTG 11-AUG-1999
DEFINITION	Plasmidium falciiparum 3D7 chromosome 13, *** SQUENCING IN PROGRESS				
	***, 9 unordered pieces.				
ACCESSION	AL109814				
VERSION	AL109814.1	GI:5731891			
KEYWORDS	HTG: HTGS PHASE1.				
SOURCE	Plasmidium falciiparum 3D7				
ORGANISM	Plasmidium falciiparum 3D7				



PFMAL13\_14  
WPCOMMENT

WPCCOMMENT

Sequence split into	28 fragments	LOCUS	PFMAL13	Accession	AL844505
Fragment Name	Pos'n	End			

Fragment Name	Begin	End
PFMAL3_00	1	110000
PFMAL3_01	100001	210000
PFMAL3_02	200001	310000
PFMAL3_03	300001	410000
PFMAL3_04	400001	510000
PFMAL3_05	500001	610000
PFMAL3_06	600001	710000
PFMAL3_07	700001	810000
PFMAL3_08	800001	910000
PFMAL3_09	900001	1010000
PFMAL3_10	1000001	1110000
PFMAL3_11	1100001	1210000
PFMAL3_12	1200001	1310000
PFMAL3_13	1300001	1410000
PFMAL3_14	1400001	1510000
PFMAL3_15	1500001	1610000
PFMAL3_16	1600001	1710000
PFMAL3_17	1700001	1810000
PFMAL3_18	1800001	1910000
PFMAL3_19	1900001	2010000
PFMAL3_20	2000001	2110000
PFMAL3_21	2100001	2210000
PFMAL3_22	2200001	2310000
PFMAL3_23	2300001	2410000
PFMAL3_24	2400001	2510000
PFMAL3_25	2500001	2610000
PFMAL3_26	2600001	2710000
PFMAL3_27	2700001	2732359

QY 277 TAATAAA 283  
Db 35378 TCTTAGA 35384

RESULT 15  
AB136154/c

LOCUS AB136154 413 bp DNA linear STS 27-OCT-2005  
DEFINITION Homo sapiens DNA, STS on chromosome 1, DIS19691, sequence tagged site.

ACCESSION AB136154  
VERSION AB136154.1 GI:62156735  
KEYWORDS STS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.

REFERENCE  
AUTHORS

1 Tamiya, G., Shinya, M., Imanishi, T., Ikuta, T., Makino, S., Okamoto, K., Furugaki, K., Matsumoto, T., Mano, S., Ando, S., Nozaki, Y., Yukawa, M., Nakashige, R., Yamaguchi, D., Ishibashi, H., Yonekura, M., Nakami, Y., Takayama, S., Endo, T., Saruwatari, T., Yagura, M., Yoshikawa, Y., Fujimoto, K., Oka, A., Chiku, S., Linsen, S. E., Giphart, M. J., Kulski, J. K., Fukazawa, T., Hashimoto, H., Kimura, M., Hoshina, Y., Suzuki, Y., Hotta, T., Mochida, J., Minezaki, T., Komai, K., Shiozawa, S., Taniguchi, A., Yamanaka, H., Kamatani, N., Gojobori, T., Bahram, S. and Inoko, H.  
Whole genome association study of rheumatoid arthritis using 27 039 microsatellites  
Hum. Mol. Genet. 14 (16), 2305-2321 (2005)

JOURNAL  
PUBMED  
REFERENCE  
AUTHORS

2 (bases 1 to 413)  
Tamiya, G., Makino, S., Fujimoto, K., Oka, A., Hayashi, H., Denda, A., Linsen, S. E., Ikuta, T., Shinya, M., Endo, T., Tomizawa, M., Tokudo, E., Sato, R., Takaki, A., Nagatsuka, Y., Watanabe, K., Adachi, S., Makino, Y., Nakano, S., Yamamoto, A., Yoshida, K., Okamoto, K., Yamaguchi, D., Ishibashi, H., Yonekura, M., Takayama, S., Nakami, Y., Saruwatari, T., Brand, A., van Hiltten, J. A., van de Watering, L. M., Giphart, M. J., Bahram, S., Kulski, J. K. and Inoko, H.  
Direct Submisison  
Submitted (12-DEC-2003) Hidetoshi Inoko, Tokai University School of Medicine, Department of Genetic Information, Bohseidai, Isehara, Kanagawa, 259-1193, Japan (E-mail: hinoko@is.icc.u-tokai.ac.jp, Tel:81-463-93-1121, Fax:81-463-94-8884)

## COMMENT

Polymorphisms were confirmed by comparing with the pooled DNA typing data of 88 Dutch population or 100 Australian. When there was no significant difference between the Japanese and these Caucasians, individual typing was performed to confirm polymorphisms.

## FEATURES

## SOURCE

Location/Qualifiers  
1. 413  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/map="1q43"

/note="pooled DNA of 100 Japanese unrelated individuals sequenced site DIS19691"  
1. 24  
/note="5' primer: AGTGTATGATATGCTTCATATGTC"  
PCR conditions="denaturation 96degC 5 min, 57degC 1 min, 72degC 1 min; 40 cycles 96degC 45 sec, 57degC 45 sec, 72degC 1 min"

## repeat\_region

/note="sequence tagged site DIS19691"

## primer\_bind

/rpt\_type="candem"  
/rpt\_unit\_seq="aat"  
complement(390..413)  
/note="3' primer: GTGGAACTGTAGTCCATTAAC"

## ORIGIN

Query Match 8.5%; Score 43.6; DB 7; Length 413;

Best Local Similarity 55.2%; Pred. No. 0.04;  
Matches 85; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY	138	TATTAAGTCTATGCAACAAAGAAATGTCATACATCTCTTACGTCATATTATTT	197
Db	256	TATAGCAACTAGCCAACTAGGCTCTATTATTATTATTATTATTATTATTATT	197
QY	198	CATTAGATAGCCGGTCTTTTACTACACTCAATATAGATGAACAGATGATGGTTACT	257
Db	196	TATTATTACTATTATTATTGCTCTCAGCTCCAAAGGCTACAGAAAGCAAGTACTAT	137
QY	258	GACTGTTTATTAAGAGGTAATTAAGATCTAT	291
Db	136	GACTTATATAGTAGAATATATGTAATCTAATTT	103

Search completed: May 28, 2006, 10:39:27  
UOD time : 2989.22 secs

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2006, 18:26:02 ; Search time 343.339 Seconds  
(without alignments)  
10397.276 Million cell updates/sec

Title: US-10-099-663-1\_COPY\_1115\_1626

Perfect score: 512  
Sequence: 1 cagtgagatccctcatccatc.....gtctgcctcacagacagaag 512

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_8:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001as:\*
- 5: Geneseq2001bs:\*
- 6: Geneseq2002as:\*
- 7: Geneseq2002bs:\*
- 8: Geneseq2003as:\*
- 9: Geneseq2003bs:\*
- 10: Geneseq2003cs:\*
- 11: Geneseq2003ds:\*
- 12: Geneseq2004as:\*
- 13: Geneseq2004bs:\*
- 14: Geneseq2005s:\*
- 15: Geneseq2006s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	512	100.0	2381	11	ADL90127 Chicken i
2	326	63.7	336	11	ADL90128 Chicken i
3	42.8	8.4	2000	11	ACL35363 Rice stre
4	42.6	8.3	2000	8	ADA71938 Rice gene
5	42.4	8.3	5908	4	AA545387 Chemical1
6	42.4	8.3	5908	6	ABK28232 DNA trans
7	41.8	8.2	10048	6	ABL70313 Chemical1
8	41.8	8.2	10048	6	AA561251 Human gen
9	40.6	7.9	7040	4	AA546439 Tumour su
10	40.6	7.9	7040	6	ABK33963 Human DNA
11	40.6	7.9	7040	8	AB210179 Haematopo
12	40.6	7.9	7040	8	AB210033 Haematopo
13	40.6	7.9	7040	8	ADA20348 Prostate
14	40.6	7.9	7040	8	ADA84155 Human ren
15	40.6	7.9	7040	10	ADBE84191 Human lym
16	40.6	7.9	7040	10	ADBE84115 Human lym
17	40.6	7.9	14095	6	ABL32477 Human imm
18	40.6	7.9	110000	14	ABD76257_3 Continuation (4 of

19	40.2	7.9	6065	6	ABK31356 Signal tr
20	40.2	7.9	6065	6	ABL70579 Chemical1
21	40.2	7.9	6065	6	AA561260 Human gen
22	40.2	7.9	8365	2	AA220056 Plasmodiu
23	39.8	7.8	2000	11	ACL35363 Rice stre
24	39.6	7.7	6086	14	ADZ64438 Human can
25	39.6	7.7	110000	14	ABE04876_4 Continuation (5 of
26	39.4	7.7	5265	6	ABK28419 DNA trans
27	39.4	7.7	110000	10	ACF67367_23 Continuation (24 of
28	39.4	7.7	110000	10	ACF65386_5 Continuation (6 of
29	39.2	7.7	9117	6	ABL33069 Human imm
30	39	7.6	2933	4	ABL18616 Human imm
31	39	7.6	5311	6	ABL33019 Human imm
32	39	7.6	6202	13	ADV35113 Human CDN
33	39	7.6	6240	10	ADD49062 Human NOV
34	39	7.6	22052	4	ABL20542 Drosophi1
35	39	7.6	26097	4	ABL03096 Drosophi1
36	39	7.6	36194	4	ABL20582 Drosophi1
37	39	7.6	59967	4	ABL15492 Drosophi1
38	38.4	7.5	32392	6	ABL56203 AMEPV gen
39	38.4	7.5	50000	6	ABL5643 AMEPV gen
40	38.2	7.5	6072	6	ABL32031 Human imm
41	38.2	7.5	56737	6	AB569895 Human hyp
42	38.2	7.5	56737	15	AEF10056 Human hyp
43	38.2	7.5	110000	12	ADH69807_4 Continuation (5 of
44	38	7.4	2000	11	ACL37108 Rice stre
45	38	7.4	6802	6	ABL33231 Human imm

## ALIGNMENTS

RESULT 1  
ADL90127 standard; DNA; 2381 BP.

ADL90127  
20-MAY-2004 (first entry)

Chicken intestinal fatty acid binding protein, tFABP, gene, 5' region.  
Chicken; ds; intestinal fatty acid binding protein; tFABP;  
gut specific promoter; transgenic.

Gallus gallus.  
US2003177516-A1.  
18-SEP-2003.

14-MAR-2002; 2002US-00099663.  
14-MAR-2002; 2002US-00099663.

(HORS/) HORSEMAN N D.  
(PRATT/) PRATT S L.  
Horseman ND, Pratt SL;  
WPI; 2003-898653/82.

New nucleic acid molecule comprising an isolated avian gut-specific gene expression control region, useful for regulating heterologous nucleic acids in transgenic avians, and for generating transgenic birds.

Claim 1; SEQ ID NO 1; 28pp; English.

The invention relates to an isolated nucleic acid comprising an isolated avian gut-specific gene expression control region appearing as ADL90127 (Chicken intestinal fatty acid binding protein, tFABP, gene, 5' region or ADL90128 (Chicken tFABP promoter) or its degenerate variant. Also included are a recombinant DNA molecule comprising an isolate avian

CC gut-specific gene expression control region operably linked to a nucleic  
CC acid insert encoding a polypeptide, an expression vector that integrates  
CC into a host cell (and comprising the isolated avian gut-specific gene  
CC expression control region), expressing a heterologous polypeptide in a  
CC host cell (by transfecting a eukaryotic cell with the recombinant DNA  
CC molecule, and culturing the transfected cell in a medium suitable for  
CC expression of a heterologous polypeptide under the control of an avian  
CC intestinal fatty acid binding protein (IFABP) or cp35 gene expression  
CC control region encoded by the recombinant DNA molecule), a eukaryotic  
CC cell transformed with the expression vector (or its progeny, which  
CC expresses a heterologous polypeptide) and a transgenic avian having a  
CC heterologous polynucleotide sequence comprising the nucleic acid insert.  
CC The nucleic acids are useful for regulating heterologous nucleic acids in  
CC transgenic avians, as probes in nucleic acid hybridisation assays for  
CC detecting the IFABP gene expression control region, and for generating  
CC transgenic birds. The present sequence is the Chicken intestinal fatty  
CC acid binding protein, IFABP, gene, 5' region.

XX Sequence 2381 BP; 775 A; 428 C; 469 G; 709 T; 0 U; 0 Other;

Query Match 100.0%; Score 512; DB 11; Length 2381;

Best Local Similarity 100.0%; Pred. No. 3.3e-129; Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGTGGATCTTCATCTCATGCTTATTTATGTGAGTAGAATGATTCACCAATTAGA 60  
DB 1115 CAGTGGATCTTCATCTCATGCTTATTTATGTGAGTAGAATGATTCACCAATTAGA 1174  
QY 61 ATGACAAAGCAGAGATTGTTGTTTATCTGTGGTAAATACGTTTCTCCAGTTGAT 120  
DB 1175 ATGACAAAGCAGAGATTGTTGTTTATCTGTGGTAAATACGTTTCTCCAGTTGAT 1234  
QY 121 AAAGACCTCCACAGATATAAGCTATGCAACAAAGAAATGTCATATCTCTT 180  
DB 1235 AAAGACCTCCACAGATATAAGCTATGCAACAAAGAAATGTCATATCTCTT 1294  
QY 181 AGTCTCATTTATTTTCTATAGATAGCGGTTTCTCACTCAATTAAGTAGAAC 240  
DB 1295 AGTCTCATTTATTTTCTATAGATAGCGGTTTCTCACTCAATTAAGTAGAAC 1354  
QY 241 AGAATGAATGGGTTATAGTACCTTTATTAAGAAAGTAAATATCATCATCTT 300  
DB 1355 AGAATGAATGGGTTATAGTACCTTTATTAAGAAAGTAAATATCATCATCTT 1414  
QY 301 AGGCAATAGGAGGAGAGAGATTGAGCAACAGTGTGCTTCAAGTGAACAAAGTTAA 360  
DB 1415 AGGCAATAGGAGGAGAGAGATTGAGCAACAGTGTGCTTCAAGTGAACAAAGTTAA 1474  
QY 361 ACTAAAGTACCCCTCTCTTGACAAGATCAATGCCACAGTTAGCTTACGCCACA 420  
DB 1475 ACTAAAGTACCCCTCTCTTGACAAGATCAATGCCACAGTTAGCTTACGCCACA 1534  
QY 421 TCATCATGTAATATGCTTCTGTAATAGCCGTTCATTAATCTCTTTGCAAACTCTGC 480  
DB 1535 TCATCATGTAATATGCTTCTGTAATAGCCGTTCATTAATCTCTTTGCAAACTCTGC 1594  
QY 481 TACTTACCAGAAAGTCTGCTACAGACAAAG 512  
DB 1595 TACTTACCAGAAAGTCTGCTACAGACAAAG 1626

RESULT 2

ADL90128 ID ADL90128 standard; DNA; 336 BP.

XX ADL90128;

XX 20-MAY-2004 (first entry)

XX Chicken intestinal fatty acid binding protein, IFABP, gene, promoter.

XX Chicken; ds; intestinal fatty acid binding protein; IFABP;

KM gut specific promoter; transgenic; promoter.

XX Gallus gallus.

XX US2003177516-A1.

XX 18-SEP-2003.

XX 14-MAR-2002; 2002US-00099663.

XX 14-MAR-2002; 2002US-00099663.

XX (HORS/) HORSEMAN N D.

XX (PRAT/) PRATT S L.

XX Horseman ND, Pratt SL;

XX WPI; 2003-898653/82.

XX New nucleic acid molecule comprising an isolated avian gut-specific gene  
XX expression control region, useful for regulating heterologous nucleic  
XX acids in transgenic avians, and for generating transgenic birds.

XX Claim 1; SEQ ID NO 2; 28bp; English.

XX The invention relates to an isolated nucleic acid comprising an isolated  
XX avian gut-specific gene expression control region appearing as  
XX ADL90127/Chicken intestinal fatty acid binding protein, IFABP, gene, 5'  
XX region or ADL90128 (chicken IFABP promoter) or its degenerate variant.  
XX Also included are a recombinant DNA molecule comprising an isolate avian  
XX gut-specific gene expression control region operably linked to a nucleic  
XX acid insert encoding a polypeptide, an expression vector that integrates  
XX into a host cell (and comprising the isolated avian gut-specific gene  
XX expression control region), expressing a heterologous polypeptide in a  
XX host cell (by transfecting a eukaryotic cell with the recombinant DNA  
XX molecule, and culturing the transfected cell in a medium suitable for  
XX expression of a heterologous polypeptide under the control of an avian  
XX intestinal fatty acid binding protein (IFABP) or cp35 gene expression  
XX control region encoded by the recombinant DNA molecule), a eukaryotic  
XX cell transformed with the expression vector (or its progeny, which  
XX expresses a heterologous polypeptide) and a transgenic avian having a  
XX heterologous polynucleotide sequence comprising the nucleic acid insert.  
XX The nucleic acids are useful for regulating heterologous nucleic acids in  
XX transgenic avians, as probes in nucleic acid hybridisation assays for  
XX detecting the IFABP gene expression control region, and for generating  
XX transgenic birds. The present sequence is the Chicken intestinal fatty  
XX acid binding protein, IFABP, gene, promoter.

XX Sequence 336 BP; 116 A; 62 C; 63 G; 95 T; 0 U; 0 Other;

Query Match 63.7%; Score 326; DB 11; Length 336;

Best Local Similarity 100.0%; Pred. No. 9.3e-79; Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 ATTATATATTTTCTATAGATAGCGGTTTTTCTACCACTCAATTAAGTAGAAGATG 246  
DB 1 ATTATATATTTTCTATAGATAGCGGTTTTTCTACCACTCAATTAAGTAGAAGATG 60  
QY 247 AATGGTTAGTACCTGTTTATTAAGAAAGTAAATATCATCATCTTGTAGGCA 306  
DB 61 AATGGTTAGTACCTGTTTATTAAGAAAGTAAATATCATCATCTTGTAGGCA 120  
QY 307 TAAGGAGGAGAGATTGACCAACAGTGTGCTTACAAGTGAACAAAGTTAACTAAA 366  
DB 121 TAAGGAGGAGAGATTGACCAACAGTGTGCTTACAAGTGAACAAAGTTAACTAAA 180  
QY 367 GTGACCCCTCTCTTGACAAGATGACAGTGTGAGCTTTAGCCACCATCATCA 426  
DB 181 GTGACCCCTCTCTTGACAAGATGACAGTGTGAGCTTTAGCCACCATCATCA 240  
QY 427 TGTAAATGCTTCTCGATTAAGCTGTTCATTAATCTCTTTGCAAAAGCTGTACTTA 486  
DB 241 TGTAAATGCTTCTCGATTAAGCTGTTCATTAATCTCTTTGCAAAAGCTGTACTTA 500

Qy 487 CCAGAGTCTGCTTACAGACAGAAAG 512  
 Db 301 CCAGAGTCTGCTTACAGACAGAAAG 326

RESULT 3  
 ID ACL35363 standard; cDNA; 2000 BP.  
 AC ACL35363;  
 XX 02-JUN-2005 (first entry)  
 XX  
 DE Rice stress-regulated promoter SEQ ID NO:13926.  
 XX  
 KM ss; abiotic stress tolerance; transgenic plant; plant; cereal;  
 KM agriculture.  
 XX  
 OS Oryza sativa.  
 XX  
 PN WO2003008540-A2.  
 XX  
 PD 30-JAN-2003.  
 XX  
 PF 21-JUN-2002; 2002WO-US019668.  
 XX  
 PR 22-JUN-2001; 2001US-0300112P.  
 XX  
 PR 24-AUG-2001; 2001US-0314662P.  
 PR 26-SEP-2001; 2001US-0325277P.  
 PR 21-NOV-2001; 2001US-0332132P.  
 XX  
 PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
 XX  
 PI Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;  
 PI Moughamer T, Provart N, Riecke D, Zhu T;  
 DR WPI; 2003-248011/24.  
 XX  
 PT New stress-responsive nucleic acid, useful for altering the  
 PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold  
 PT stress, salt stress or osmotic stress.  
 XX  
 PS Claim 48; SEQ ID NO 13926; 89pp; English.  
 XX  
 CC The invention relates to novel abiotic stress responsive polynucleotides  
 CC and polypeptides. Also disclosed are vectors, expression cassettes, host  
 CC cells, and plants containing such polynucleotides. Also disclosed are  
 CC methods for using the polynucleotides and polypeptides to alter the  
 CC responsiveness of a plant to abiotic stress. The invention is useful in  
 CC agriculture. The nucleic acid is useful for determining whether a test  
 CC plant has been exposed to an abiotic stress condition. It is also useful  
 CC for selecting an agent that alters abiotic stress regulated  
 CC polynucleotide expression in a plant cell, and to identify a homolog or  
 CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid  
 CC molecule and the polypeptide encoded by it are useful in altering the  
 CC responsiveness of a plant to an abiotic stress, such as cold stress, salt  
 CC stress, osmotic stress or any of their combinations. The present sequence  
 CC is used in the exemplification of the invention  
 XX  
 SQ Sequence 2000 BP; 540 A; 344 C; 416 G; 428 T; 0 U; 272 Other;  
 XX

Query Match 8.4%; Score 42.8; DB 11; Length 2000;  
 Best Local Similarity 11.7%; Pred. No. 0.3; Indels 0; Gaps 0;  
 Matches 38; Conservative 148; Mismatches 140;

Qy 14 ATCTCATGCTTATATGAGTAGATGATCTCACCAATTAGATGACAAAGAG 73  
 Db 19 WKYMMMKMKKAYMYKGGAKWMTWSAVMAYAMVAKKWRMTMAKRCSCSKVMSMA 78  
 Qy 74 AGATTGTTGTTTATCTGTTGGTAATAAGCTTTCTCAAGTTGATATAAGACCTCCCA 133  
 Db 79 SKRMKMMWMAKMGSKMTYMGKRMRYKAYMKRMWMAWMSWTRATYMMWKACTYCYT 138

Qy 134 CCAGATTAAGTCTTATGCAACAAAGAAATGCTAATCATCTTCTTATGATTA 193  
 Db 139 MYKMYWYRMAAAMRWMMYKMKWABGWSAKOKAVRYRKKCTYYAMAMAMMTTTCMC 198

Qy 194 TTTTCATTAGATGCGCGTTTCTTACTCAACTCAATTAAGATGAACGAATGCGT 253  
 Db 199 TWSWYAMWTGYATWMAWMTATTTTTRMAWMTTRAMWMTWRYAMWYTWMAKMMKAT 258

Qy 254 TAGTCACTGTTTATTAAGAAAGTAAATGATATCATCATCTTGAAGCAATAAGGA 313  
 Db 259 KMTTYAMRWMMWMMMSRTTYRMWMAWMTWSWYTAAMSWMWCMWTTMSRRKMYR 318

Qy 314 GAGAGATTCAGCAACAGTGTCT 339  
 Db 319 YKSRKRTWTWTWTWMMAMKMKYK 344

RESULT 4  
 ID ADA71938/c  
 ID ADA71938 standard; DNA; 2000 BP.  
 AC ADA71938;  
 XX  
 DT 20-NOV-2003 (first entry)  
 XX  
 DE Rice gene, SEQ ID 5263.  
 XX  
 KM Plant; bacterial infection; fungal infection; viral infection; rice;  
 KM gene; ds.  
 XX  
 OS Oryza sativa.  
 XX  
 PN WO200300898-A1.  
 XX  
 PD 03-JAN-2003.  
 XX  
 PF 22-JUN-2001; 2001WO-IB001105.  
 XX  
 PR 22-JUN-2001; 2001WO-IB001105.  
 XX  
 PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
 XX  
 PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
 PI Katagiri F, Quan Y, Tao Y, Whitman S, Xie Z, Zhu T, Zou G;  
 XX  
 DR WPI; 2003-175290/17.  
 XX  
 PT Identifying at least one gene involved in plant resistance or response to  
 PT pathogenic infection for conferring resistance or tolerance to a plant to  
 PT bacterial, fungal or viral infection by determining or detecting plant  
 PT gene expression.  
 XX  
 PS Claim 27; SEQ ID NO 5263; 89pp; English.  
 XX  
 CC The present invention relates to a method (M1) for identifying genes  
 CC involved in plant resistance or response to pathogenic infection. M1  
 CC comprises identifying a gene whose expression is significantly altered in  
 CC the incompatible interaction of plant gene expression relative to  
 CC expression of the gene in an uninfected plant, in a mutant plant that  
 CC does not express a gene associated with response to pathogenic infection,  
 CC or in a corresponding incompatible or compatible interaction. (M1) is  
 CC useful for conferring resistance to resistance or tolerance to a plant to  
 CC bacterial, fungal or viral infection. The present sequence was used to  
 CC illustrate the invention.  
 XX  
 SQ Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;  
 XX

Query Match 8.3%; Score 42.6; DB 8; Length 2000;  
 Best Local Similarity 10.6%; Pred. No. 0.34; Indels 1; Gaps 1;  
 Matches 44; Conservative 194; Mismatches 178;

Qy 1 CAGTGATCCTTCATCTCATGCTTATATGAGTAGATGATTCACCAATTAGA 60

```

Db      1016 MRRYTRSMSTWYAMMKYTKYMTAYSSTWKMYAYKRAYAMRSRKTWCTGCKMAT 957
Qy      61 ATGCAAGCAAGCAGACATTTGTTTATCTGTTGGTAATAAGTTTTCACAGTTGAT 120
Db      956 YCGTMMAGKMRMAMCMYCCMMKMKMTSCMMWKYRWSCMYTMMGAMRAYAM 897
Qy      121 AAAGACCTCCACCAAGTAAAGTCTATGCAACAAGAA-AATGTCATATCTTCT 179
Db      896 RRRRTYKMSWRMYMTWTKMAWMTMCMAKRYMATGATMMMYTTCYATCA 837
Qy      180 TAGCTCATTTATTTTATTTATATAGACCGGTTTCTTACTCAACTCAATAAGTGA 239
Db      836 KCKTKMNTKMTWTWACAMRATSWMRAMKRYMRYKRAYMRWWRMCKAGNARWK 777
Qy      240 CAGAATGAATGGTTAGTACTGTTTATAAGAAAGATAATAAGATCATCATTT 299
Db      776 SRYMKMKKATATRYMMOMAMTMMWSRMRKSTRMMSGMRMWSAMRYCSRCKATRY 717
Qy      300 GAGCCATAAGGAGGAGAGATTCAGCAAAACAGTGTCTTACAGTGAAGAAACAATT 359
Db      716 ASSARMTKRAKRSYRRRRYMKRGWTYRYRYSRSCMTRAMSRKRRMAGSMKSCM 657
Qy      360 AACTAAAGTACCCCTCTGCAAGATCATGCAAGTTCAGTTCAGTTCAGTTCAGTTC 416
Db      656 YMRGASMMWYSKXSCAKCKKTRMTSTMTGMTSYSSKSMSTSKMSYMGKM 600

RESULT 5
AAS45387/c
ID      AAS45387 standard; DNA; 5908 BP.
AC
AA
AS45387;
XX
DT      18-DEC-2001 (first entry)
XX
DE      Chemically pretreated complementary DNA associated with cell cycle #46.
XX
KW      Cell cycle; human; CpG dinucleotide; cytosine methylation; HIV; aging;
KW      human immunodeficiency virus; neurodegenerative disorder; solid tumour;
KW      graft-versus-host disease; glomerular disease; Lewy body disease; cancer;
KW      arthritis; arteriosclerosis; anti-HIV; neuroprotective; antiarthritic;
KW      immunosuppressive; antitumour; cytosolic; antiarteriosclerotic; ds;
KW      PCR primer.
XX
OS      Homo sapiens.
XX
PN      WO200168911-A2.
XX
PD      20-SEP-2001.
XX
PF      15-MAR-2001; 2001WO-EP002945.
XX
PR      06-MAR-2000; 2000DE-01013847.
PR      06-APR-2000; 2000DE-01019058.
PR      07-APR-2000; 2000DE-01019173.
PR      30-JUN-2000; 2000DE-01032529.
PR      01-SEP-2000; 2000DE-01043826.
XX
PA      (EPIG-) EPIGENOMICS AG.
XX
PI      Olek A, Piepenbrock C, Berlin K;
XX
DR      WPI; 2001-602751/68.
XX
PT      Designing primers and probes for analyzing diseases associated with
PT      cytosine methylation state e.g. arthritis, cancer, aging,
PT      arteriosclerosis comprising fragments of chemically modified genes
XX
XX      associated with cell cycle.
XX
PS      Claim 1, SEQ ID NO 92; 28bp; English.
CC      Sequences AAS45296-AAS45520 represent chemically pretreated genomic DNA
CC      molecules associated with the cell cycle and specific PCR primers of the

```

```

CC      invention. The sequences are useful for detecting the methylation state
CC      of all CpG dinucleotides in a sequence and therefore for analysing
CC      associated diseases. By analysing cytosine methylations in the pretreated
CC      DNA, genetic and/or epigenetic parameters for the diagnosis and therapy
CC      of existing diseases or the predisposition to specific diseases can be
CC      ascertained. The parameters may be compared to another set of genetic
CC      and/or epigenetic parameters, the differences serving as basis for
CC      diagnosis and/or prognosis events which are disadvantageous to patients.
CC      The sequences of the invention are useful for the diagnosis and therapy
CC      of HIV infection, neurodegenerative disorders, graft-versus-host disease,
CC      aging, glomerular disease, Lewy body disease, arthritis,
CC      arteriosclerosis, solid tumours and cancers
SQ      Sequence 5908 BP; 1706 A; 149 C; 1421 G; 2632 T; 0 U; 0 Other;
Query Match      8.3%; Score 42.4; DB 4; Length 5908;
Best Local Similarity 54.5%; Pred. No. 0.55;
Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
Qy      138 TATTAAGTCTTATGCAAGAAAGAAATGTCATATCATTTCTTATGCTATTTATTT 197
Db      4703 TATAACTCTTTTATATTAACAAATATATATATATCTATTAACCATCTTTTAA 4644
Qy      198 CATTAAGTACCGGTTTCTTACTCAACTCAATAAGATGAACAGATGAATGGTTACT 257
Db      4643 TATTATATTTTATTTTCACTACCTTCTCTTAAATTTAAAAATTAATCTATACCTTAA 4584
Qy      258 GACTGTTTATTAAGAAAGATTAATAAGATATATCA 293
Db      4583 TCCCTTTATCCAAATTAATTAATCTATATACGAACA 4548

RESULT 6
ABK28232/c
ID      ABK28232 standard; DNA; 5908 BP.
XX
AC      ABK28232;
XX
DT      23-APR-2002 (first entry)
XX
DE      DNA transcription associated complementary genomic DNA #53.
XX
KW      DNA transcription associated gene; peptide nucleic acid; pNA-oligomer;
KW      pNA; cytosine methylation state; SNP; retroviral infection; gene; ds;
KW      single nucleotide polymorphism; adenovirus deamase deficiency; cancer;
KW      viral infection; Sezary syndrome; haematological disorder; tuberculosis;
KW      immunological disorder; Werner syndrome; developmental disorder;
KW      psoriasis; Rieger's syndrome; neurological disorder; erythropoiesis;
KW      neurodegenerative disorder; Maardenburg syndrome; Niemann-Pick disease;
KW      myelodysplastic syndrome; myocardial infarction; hypertension; arthritis;
KW      angiodysgenesis; congenital heart disease; HDR syndrome; gene therapy;
KW      polyglutamine disorder; solid tumour.
XX
OS      Unidentified.
XX
PN      WO200192565-A2.
XX
PD      06-DEC-2001.
XX
PF      06-APR-2001; 2001WO-EP003973.
XX
PR      06-APR-2000; 2000DE-01019058.
PR      07-APR-2000; 2000DE-01019173.
PR      30-JUN-2000; 2000DE-01032529.
PR      01-SEP-2000; 2000DE-01043826.
XX
PA      (EPIG-) EPIGENOMICS AG.
XX
PI      Olek A, Piepenbrock C, Berlin K;
XX
DR      WPI; 2002-090046/12.
XX
PT      New nucleic acids or oligomers, useful for diagnosing or treating

```



OS Homo sapiens.  
XX WO200177375-A2.  
XX 18-OCT-2001.  
XX  
XX 06-APR-2001; 2001WO-EP003968.  
XX  
XX 06-APR-2000; 2000DE-01019058.  
PR 07-APR-2000; 2000DE-01019173.  
PR 30-JUN-2000; 2000DE-01032529.  
PR 01-SEP-2000; 2000DE-01043826.  
XX  
XX (EPIG-) EPIGENOMICS AG.  
XX  
XX Olek A, Piepenbrock C, Berlin K;  
XX WPI, 2002-017470/02.  
XX  
XX New nucleic acid sequences from chemically modified genes associated with  
PT gene regulation, useful for analyzing cytosine methylations for diagnosis  
PT and therapy of diseases e.g. severe combined immunodeficiency disease.  
XX  
XX Claim 1; SEQ ID NO 211; 26bp; English.  
XX  
XX The invention relates to 224 nucleic acid sequences comprising at least  
CC 18 bases of a chemically pretreated gene associated with gene regulation  
CC selected from 43 known genes (or complementary sequences). The chemical  
CC pretreatment converts cytosine bases unmethylated at the 5-position to  
CC uracil or another base with hybridisation behaviour dissimilar to  
CC cytosine, to enable analysis of cytosine methylations. The DNA sequences,  
CC oligomers (or sets/arrays) and method are useful in the diagnosis of  
CC diseases (or predisposition to diseases) associated with gene regulation  
CC and in therapy of such diseases, by enabling analysis of the cytosine  
CC methylation patterns of such genes, kits are provided. They are  
CC especially useful in diagnosis and therapy of e.g. severe combined  
CC immunodeficiency disease, cardiac disorders, haemophilia, solid tumours  
CC and cancer, Werner syndrome, asthma, HDR syndrome, Saethre-Chotzen  
CC syndrome, renal disease, preclampsia, graft versus-host disease. The  
CC present sequence is a sequence included in the sequence data for this  
CC specification and is associated with the human gene regulation-associated  
CC genes. Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_sequences  
XX  
XX Sequence 10048 BP; 2873 A; 57 C; 1888 G; 5225 T; 0 U; 5 Other;  
SQ  
Query Match 8.2%; Score 41.8; DB 6; Length 10048;  
Best Local Similarity 46.4%; Pred. No. 0.95;  
Matches 136; Conservative 0; Mismatches 157; Indels 0; Gaps 0;  
OY 100 ATACGTTTCTCCAGTTGTAATAAGACCTCCACAGTATTAAGTCTATGCAACAAG 159  
DB 5471 ATTTTATTTATCTTAAACTTAATATTCCTCTCAAAAAAAATTAATATCTATAC 5412  
OY 160 AAAATGTCATACATCTCTTCTAGTCTCATTTATTTTCATTAAGTAGCGGTTTTTAC 219  
DB 5411 TCCAAATCCAAAAATCTATTTTAAACCTATTAACCTCTCAAAATATCTTCTAC 5352  
OY 220 TACAACCTAAATTAAGATGAACGAATGAATGGGTTAGTACTGTTTAAAGAGAGTAA 279  
DB 5351 AAATTACTTAATTAATAAACACACACTTAATCTTAACTTAATCAAAATACAAAAAAC 5292  
OY 280 TAAAGATACATATCATTTGAGGCAATTAAGGAGGAGAGATTCACCAACAAGTGTCT 339  
DB 5291 TATTAATTTACTTTTAAATTAATTAATTAATTAATTTTAAATTAATTAATTAATTTTCA 5232  
OY 340 TACAAGTGAAGAAACAGTTAAAGTGAACCCCTCTCTTGAACAAGATCAA 392  
DB 5231 AAACAAAACAAAACAAAACAAAACAAAACCAACCAATTTTCAAACTATATCTA 5179

5 RESULT 9

AA546439/c  
ID AA546439 standard; DNA; 7040 BP.  
XX  
XX AC AA546439;  
XX  
XX 18-DEC-2001 (first entry)  
XX  
XX Tumour suppressor gene derived chemically modified sequence #161.  
XX Human; tumour suppressor gene; oncogene; antitumour; cytosine; cancer;  
XX tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;  
XX cytosine methylation; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO200168912-A2.  
XX  
XX 20-SEP-2001.  
XX  
XX 15-MAR-2001; 2001WO-EP002955.  
XX  
XX 15-MAR-2000; 2000DE-01013847.  
PR 06-APR-2000; 2000DE-01019058.  
PR 07-APR-2000; 2000DE-01019173.  
PR 30-JUN-2000; 2000DE-01032529.  
PR 01-SEP-2000; 2000DE-01043826.  
XX  
XX (EPIG-) EPIGENOMICS AG.  
XX  
XX Olek A, Piepenbrock C, Berlin K;  
XX WPI, 2001-602752/68.  
XX  
XX Fragments of chemically modified genes associated with tumor suppressor  
PT genes and oncogenes, useful in designing primers and probes for analyzing  
PT diseases associated with cytosine methylation state e.g. cancer.  
XX  
XX Claim 1; SEQ ID NO 161; 27bp; English.  
XX  
XX The invention relates to a nucleic acid comprising a sequence of 18  
XX bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with  
XX bisulphite, of genes associated with tumor suppression and oncogenes  
XX having a sequence taken from 536 (actually 533 since numbers 408, 458 and  
XX 500 are missing from the sequence listing) sequences (Ss) and sequences  
XX complementary to (Ss). The nucleic acid may be a peptide nucleic acid-  
XX oligomer (PNA) of at least 9 nucleotides and may form part of a set of  
XX probes for detecting the cytosine methylation state and/or single  
XX nucleotide polymorphisms and also to be used in an array for analysing  
XX diseases associated with CpG dinucleotides e.g. cancers and tumours. The  
XX probes can also be used in a method for ascertaining genetic and/or  
XX epigenetic parameters for the diagnosis and/or therapy of existing  
XX diseases or the predisposition to specific diseases, by analysing  
XX cytosine methylations. The parameters may be compared to another set of  
XX genetic and/or epigenetic parameters, the differences serving as basis  
XX for diagnosis and/or prognosis events which are disadvantageous to  
XX patients. The present sequence is one of the 533 genomic sequences  
XX derived from tumor suppressor genes and oncogenes. Note: The sequence  
XX data for this patent did not form part of the printed specification, but  
XX was obtained in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_sequences  
SQ  
Sequence 7040 BP; 2018 A; 184 C; 1566 G; 3272 T; 0 U; 0 Other;  
Query Match 7.9%; Score 40.6; DB 4; Length 7040;  
Best Local Similarity 50.8%; Pred. No. 1.8;  
Matches 97; Conservative 0; Mismatches 94; Indels 0; Gaps 0;  
OY 101 TACGTTTCTCCAGTTGTAATAAGACCTCCACAGTATTAAGTCTATGCAACAAG 160  
DB 3586 TATATTTTAAAAATCAAAATTAATATTCCTCTTCAAAACATTAATTAATTT 3527  
OY 161 AAATGTCATACATCTCTTCTAGTCTCATTTATTTTCAATTAATGCGGTTTTTACT 220

Db 3526 ATATATTAACCAATTTTAAATTAACAAATTCATTTTAAATATATTAACCTTAAAAA 3467  
Oy 221 ACAACTCAAAATAGATGAGACAGATGATGGTTAGTGACTGTTTAAAGAAGACTAAT 280  
Db 3466 ATTAATTTAATAACATTAACCAATTTTCAAAAAACAATTTAAATCTATTAACCTTAAT 3407  
Oy 281 AAAGATTAAT 291  
Db 3406 AAAATTTCTAT 3396

RESULT 10  
ABK33963/c  
ID ABK33963 standard; DNA; 7040 BP.  
XX  
XX ABK33963;  
XX  
XX 18-JUN-2002 (first entry)  
XX  
XX Human DNA for staging of Astrocytomas #24.  
XX  
XX Human; ds; astrocytoma; cytosstatic; staging; cysteine methylation; CpG;  
XX bisulphite; brain tissue; MALDI; ESI; electron spray mass spectrometry;  
XX matrix assisted laser desorption/ionization mass spectrometry.  
XX  
XX Homo sapiens.  
XX  
XX WO200202808-A2.  
XX  
XX 10-JAN-2002.  
XX  
XX 02-JUL-2001; 2001WO-EP007538.  
XX  
XX 30-JUN-2000; 2000DE-01032529.  
XX  
XX 01-SEP-2000; 2000DE-01043826.  
XX  
XX (EPIG-) EPIGENOMICS AG.  
XX  
XX Olek A, Piepenbrock C, Berlin K;  
XX  
XX WPI; 2002-171649/22.  
XX  
XX Novel chemically modified genomic DNA sequences, useful in the  
XX characterization, classification, differentiation, grading, staging,  
XX treatment and/or diagnosis of astrocytomas or predisposition to  
XX astrocytomas.  
XX  
XX Claim 1; SEQ ID NO 47; 37pp; English.  
XX  
XX The invention relates to a nucleic acid comprising a sequence (I) of at  
XX least 18 bases in length of a segment of chemically pre-treated genomic  
XX DNA which has any one of the sequences of (ABK33913-ABK34032) or its  
XX complement. Also included are an oligonucleotide or peptide nucleic acid  
XX (or set thereof) of at least 9 nucleotides which hybridises to (I),  
XX primers for (I), probes for detecting cytosine methylation or single-  
XX nucleotide polymorphisms (SNP) in (I), an array of oligomers or peptide  
XX nucleic acids for analysing diseases associated with the methylation  
XX states of the CpG dinucleotides of (I). The array is useful for  
XX determining genetic and/or epigenetic parameters, classification,  
XX differentiation, grading, staging, treatment and/or diagnosis of  
XX astrocytomas, or the predisposition to astrocytomas by analysing cytosine  
XX methylations, involves obtaining a biological sample containing genomic  
XX DNA, extracting the genomic DNA, converting cytosine bases which are  
XX unmethylated at the 5-position, in the genomic DNA sample, to uracil or  
XX another base which is dissimilar to cytosine in terms of hybridisation  
XX behaviour, by chemical treatment and amplifying chemically pre-treated  
XX genomic DNA fragments using the array and a polymerase, where the  
XX amplicates carry a detectable label. The method further involves  
XX identifying methylation status of one or more cytosine positions, and  
XX analysing methylation status of the cytosine positions by reference to  
XX one or more data sets. The genomic DNA is chemically treated by using a  
XX bisulphite, hydrogen sulphite or disulphite. The amplification step  
XX amplifies DNA which is of particular interest in astrocytoma or brain

CC tissue, based on the specific genomic methylation status of brain  
CC tissues, as opposed to background DNA. The amplicates carry a  
CC fluorescent label or radionuclide. Optionally, the labels of the  
CC amplicates are detachable molecule fragments having a typical mass  
CC which are detected in a mass spectrometer. The fragments of chemically  
CC pre-treated genomic DNA to be amplified, have a single positive or  
CC negative charge for a better detectability in the mass spectrometer.  
CC Preferably, the amplicates or fragments of the amplicates are  
CC detected by matrix assisted laser desorption/ionization mass spectrometry  
CC (MALDI) or using electron spray mass spectrometry (ESI). The present  
CC sequence is one of the chemically pre-treated reference DNA samples of  
CC the invention. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
XX directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences)

SO Sequence 7040 BP; 2018 A; 184 C; 1566 G; 3272 T; 0 U; 0 Other:  
Query Match 7.9%; Score 40.6; DB 6; Length 7040;  
Best Local Similarity 50.8%; Pred. No. 1.8;  
Matches 97; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

Oy 101 TACGTTTCTCCAGTTGATTAAGACCCCTCCACAGTATAAGCTTATGACAAACA 160  
Db 3586 TATATTTAAAAAACAATTAATAAATTCCTTTCTAAAAACAATTATATTAAT 3527  
Oy 161 AAATGCAATACATCTCTAGTCTCATATATTTTCATAGATGCGGTTTTTACT 220  
Db 3526 ATATATTAACCAATTTTAAATTAACAAATTCATTTTAAATTAATTAATTA 3467  
Oy 221 ACAACTCAAAATAGATGAGACAGATGATGGTTAGTGACTGTTTAAAGAAGACTAAT 280  
Db 3466 ATTAATTTAATAACATTAACCAATTTTCAAAAAACAATTTAAATCTATTAACCTTAAT 3407  
Oy 281 AAAGATTAAT 291  
Db 3406 AAAATTTCTAT 3396

RESULT 11  
AB210179/c  
ID AB210179 standard; DNA; 7040 BP.  
XX  
XX AB210179;  
XX  
XX 16-JAN-2003 (first entry)  
XX  
XX Haematopoietic cell proliferation disorder related DNA sequence #319.  
XX  
XX Human; haematopoietic cell proliferation disorder; cytosstatic;  
XX gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;  
XX cytosine methylation state; gene; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO20027272-A2.  
XX  
XX 03-OCT-2002.  
XX  
XX 26-MAR-2002; 2002WO-EP003401.  
XX  
XX 26-MAR-2001; 2001US-0278333P.  
XX  
XX (EPIG-) EPIGENOMICS AG.  
XX  
XX Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;  
XX Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu B;  
XX Lewin A, Lipscher B, Maier S, Model F, Mueller V, Otto T, Pelet C;  
XX Schwope I, Ziebrich H;  
XX  
XX WPI; 2003-018942/01.  
XX  
XX Detecting and differentiating between hematopoietic cell proliferative  
XX disorders, comprises contacting a target nucleic acid with a reagent that

PT distinguishes between methylated and non-methylated CpG dinucleotides.  
 XX  
 XX Claim 28; SEQ ID NO 319; 117bp; English.  
 PS  
 CC The present invention describes a method for detecting and  
 CC differentiating between haematopoietic cell proliferative disorders  
 CC associated with at least 1 gene and/or their regulatory regions in a  
 CC subject. The method comprises contacting a target nucleic acid in a  
 CC biological sample obtained from the subject with at least 1 reagent,  
 CC which distinguishes between methylated and non-methylated CpG  
 CC dinucleotides within the target nucleic acid. AB209861 to AB211118  
 CC represent specifically claimed nucleotide sequences from the present  
 CC invention. Oligonucleotides from the present invention can be used: for  
 CC differentiating between healthy haematopoietic cells and proliferative  
 CC disorder haematopoietic cells; for differentiating between acute  
 CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for  
 CC determining the cytosine methylation state and/or single nucleotide  
 CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder  
 CC related sequences and their complements; and as primers for the  
 CC amplification of haematopoietic cell proliferation disorder related DNA  
 CC sequences. The nucleotide sequences from the present invention can also  
 CC be used for detecting a predisposition to, differentiation between  
 CC subclasses, diagnosis, prognosis, treatment and/or monitoring of  
 CC haematopoietic cell proliferative disorders. The present method enables a  
 CC highly specific classification of haematopoietic cell proliferative  
 CC disorders allowing for improved and informed treatment of patients  
 CC  
 SQ Sequence 7040 BP; 2018 A; 0 C; 1566 G; 3456 T; 0 U; 0 Other;  
 XX  
 XX Query Match 7.9%; Score 40.6; DB 8; Length 7040;  
 XX Best Local Similarity 50.8%; Pred. No. 1.8;  
 XX Matches 97; Conservative 0; Mismatches 94; Indels 0; Gaps 0;  
 QY 101 TACGTTTCTCCAGTGTATTAAGACCTCCACAGTATTAAGTCTATGCAACAAGA 160  
 DB TATATTTTAAAAAATCAATATATTAATCCCTTCTTCAAAAACCTATATATTAATTT 3527  
 QY 161 AAATGTCATACATCTCTTACTGCTCATTTATTTTCAATGATAGCCGGTTTTTACT 220  
 DB ATATATTAACCACTTTTATATTAATCAATTTATTTTAAATATTAATTAATTTAAAAA 3467  
 QY 221 ACAACTCAAAATTAAGTAAGATGAATGAGTTAGTGACTGTTTATTAAGAAGATAT 280  
 DB ATTAATTAATTAACATTAACATTTTACAAAAAACATTTTAAATCTATTAATAACCTTAAT 3407  
 QY 281 AAAGATACTAT 291  
 DB 3406 AAAATTCTAT 3396  
 XX  
 XX RESULT 12  
 XX AB210033/C  
 XX ID AB210033 standard; DNA; 7040 BP.  
 XX AC  
 XX AB210033;  
 XX DT 16-JAN-2003 (first entry)  
 XX DE Haematopoietic cell proliferation disorder related DNA sequence #173.  
 XX KW Human; haematopoietic cell proliferation disorder; cytostatic;  
 XX gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia,  
 XX cytosine methylation state; gene; ds.  
 XX OS Homo sapiens.  
 XX PN WO20027272-A2.  
 XX PD 03-OCT-2002.  
 XX PF 26-MAR-2002; 2002WO-EP003401.  
 XX PR 26-MAR-2001; 2001US-027833P.

XX  
 PA (EPIC-) EPIGENOMICS AG.  
 XX  
 PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;  
 PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Leech R, Leu E, Pellet C;  
 PI Levin A, Lipscher B, Mater S, Model F, Mueller V, Otto T, Pelet C;  
 PI Schwobe I, Ziebarth H;  
 DR WPI; 2003-018942/01.  
 XX  
 XX Detecting and differentiating between hematopoietic cell proliferative  
 XX disorders, comprises contacting a target nucleic acid with a reagent that  
 XX distinguishes between methylated and non-methylated CpG dinucleotides.  
 PS Claim 28; SEQ ID NO 173; 117bp; English.  
 XX  
 XX The present invention describes a method for detecting and  
 XX differentiating between haematopoietic cell proliferative disorders  
 XX associated with at least 1 gene and/or their regulatory regions in a  
 XX subject. The method comprises contacting a target nucleic acid in a  
 XX biological sample obtained from the subject with at least 1 reagent,  
 XX which distinguishes between methylated and non-methylated CpG  
 XX dinucleotides within the target nucleic acid. AB209861 to AB211118  
 XX represent specifically claimed nucleotide sequences from the present  
 XX invention. Oligonucleotides from the present invention can be used: for  
 XX differentiating between healthy haematopoietic cells and proliferative  
 XX disorder haematopoietic cells; for differentiating between acute  
 XX lymphocytic leukaemia and acute myelogenous leukaemia; as probes for  
 XX determining the cytosine methylation state and/or single nucleotide  
 XX polymorphisms (SNPs) of haematopoietic cell proliferation disorder  
 XX related sequences and their complements; and as primers for the  
 XX amplification of haematopoietic cell proliferation disorder related DNA  
 XX sequences. The nucleotide sequences from the present invention can also  
 XX be used for detecting a predisposition to, differentiation between  
 XX subclasses, diagnosis, prognosis, treatment and/or monitoring of  
 XX haematopoietic cell proliferative disorders. The present method enables a  
 XX highly specific classification of haematopoietic cell proliferative  
 XX disorders allowing for improved and informed treatment of patients  
 XX  
 SQ Sequence 7040 BP; 2018 A; 184 C; 1566 G; 3272 T; 0 U; 0 Other;  
 XX  
 XX Query Match 7.9%; Score 40.6; DB 8; Length 7040;  
 XX Best Local Similarity 50.8%; Pred. No. 1.8;  
 XX Matches 97; Conservative 0; Mismatches 94; Indels 0; Gaps 0;  
 QY 101 TACGTTTCTCCAGTGTATTAAGACCTCCACAGTATTAAGTCTATGCAACAAGA 160  
 DB TATATTTTAAAAAATCAATATATTAATCCCTTCTTCAAAAACCTATATATTAATTT 3527  
 QY 161 AAATGTCATACATCTCTTACTGCTCATTTATTTTCAATGATAGCCGGTTTTTACT 220  
 DB ATATATTAACCACTTTTATATTAATCAATTTATTTTAAATATTAATTAATTTAAAAA 3467  
 QY 221 ACAACTCAAAATTAAGTAAGATGAATGAGTTAGTGACTGTTTATTAAGAAGATAT 280  
 DB ATTAATTAATTAACATTAACATTTTACAAAAAACATTTTAAATCTATTAATAACCTTAAT 3407  
 QY 281 AAAGATACTAT 291  
 DB 3406 AAAATTCTAT 3396  
 XX  
 XX RESULT 13  
 XX ADA20348/C  
 XX ID ADA20348 standard; DNA; 7040 BP.  
 XX AC  
 XX ADA20348;  
 XX DT 20-NOV-2003 (first entry)  
 XX DE Prostate tumour related genomic DNA sample #7.  
 XX KW cytostatic; gene therapy; genetic marker; epigenetic parameter;



KW classification; differentiation; diagnosis; prostate tumour;  
KW prostate cancer; cytosine methylation; uracil;  
KW single nucleotide polymorphism; SNP; prostate carcinoma; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO2002103042-A2.  
XX  
PD 27-DEC-2002.  
XX  
PP 14-JUN-2002; 2002WO-BP006605.  
XX  
PR 14-JUN-2001; 2001DE-01028508.  
XX  
PA (EPIC-) EPIGENOMICS AG.  
XX  
PI Dietler J, Model F, Adorjan P;  
XX  
DR WPI, 2003-167536/16.  
XX  
PT Determining genetic and/or epigenetic parameters, useful for the  
PT classification, differentiation and/or diagnosis of prostate tumors or a  
PT predisposition to prostate cancer, comprises analyzing cytosine  
PT methylation.  
XX  
PS Claim 28; Page 68-72; 376pp; English.  
XX  
CC The invention relates to a method of determining genetic and/or  
CC epigenetic parameters for the classification, differentiation and/or  
CC diagnosis of prostate tumors or the predisposition to prostate cancer,  
CC by analyzing cytosine methylation in a sample of genomic DNA. The method  
CC comprises chemically treating unmethylated cytosine bases at the 5-  
CC position to uracil or another base, which is dissimilar to cytosine in  
CC terms of hybridization behaviour; followed by amplifying at least one  
CC fragment of the chemically pre-treated genomic DNA using sets of primer  
CC oligonucleotides and a polymerase. The oligomers or probes derived from  
CC them are useful for detecting the methylation state of all CpG  
CC dinucleotides and/or single nucleotide polymorphisms (SNPs) in a  
CC chemically pre-treated genomic DNA. They are all useful for treating  
CC prostate carcinoma. This sequence represents a fragment of genomic DNA  
CC used in the method of the invention.  
XX  
SQ Sequence 7040 BP; 2018 A; 184 C; 1566 G; 3272 T; 0 U; 0 Other;  
Query Match 7.9%; Score 40.6; DB 8; Length 7040;  
Best Local Similarity 50.8%; Pred. No. 1.8; Mismatches 0; Gaps 0;  
Matches 97; Conservative 0; Indels 94; Indels 0; Gaps 0;  
QY 101 TACGTTTCTCCAGTTGTATTAAGACCTCCACAGATATAAGTCTATGCAACAAAGA 160  
DB 3586 TATATTTTAAAAAATCAATATTAATAAATCCCTTCTCTAATAACCTATATATTAATT 3527  
QY 161 AAATGTCATACATTCTCTTACTGCTCATTTATTTTTCATTAGATAGCCGGTTTTTACT 220  
DB 3526 ATATATAACCACTTTTATATATAACAATTCATTTTAAAAATATATTAACCTTTAAAAA 3467  
QY 221 ACAACTCAAAATTAAGATGAACGAATGATGGTTAGTACGTTTATTAAGAAGGTAT 280  
DB 3466 ATTAATTTAATTAACATTAACCAATTACAAAAACAATTTAAATCTATATAAACCTTAAAT 3407  
QY 281 AAAGTACTAT 291  
DB 3406 AAAATTTCTAT 3396  
RESULT 14  
ID ADA84155 standard; DNA; 7040 BP.  
XX  
AC ADA84155;  
XX  
DT 20-NOV-2003 (first entry)  
XX

DE Human renal/prostate carcinoma associated DNA SEQ ID NO:13.  
XX ds; renal cancer; prostate cancer; cytosine methylation;  
KW single nucleotide polymorphism; histological; cytological.  
XX  
OS Homo sapiens.  
XX  
PN WO2002103041-A2.  
XX  
PD 27-DEC-2002.  
XX  
PP 14-JUN-2002; 2002WO-BP006603.  
XX  
PR 14-JUN-2001; 2001DE-01028509.  
XX  
PA (EPIC-) EPIGENOMICS AG.  
XX  
PI Dietler J, Model F, Adorjan P;  
XX  
DR WPI, 2003-183991/18.  
XX  
PT Method for characterizing, classifying and/or differentiating renal and  
PT prostate cancers, by analyzing the genetic and/or epigenetic parameters  
PT of genomic DNA, particularly by determining its cytosine methylation  
PT status.  
XX  
PS Claim 1; Page 50-51; 211pp; English.  
XX  
CC The invention relates to a novel method for characterizing, classifying  
CC and/or differentiating renal and prostate cancer. The method comprises  
CC extracting genomic DNA from a biological sample, converting cytosine  
CC bases (by chemical treatment) that are unmethylated at the 5-position to  
CC uracil or another base, and amplifying at least one fragment of the  
CC chemically pretreated genomic DNA using sets of primer oligonucleotides  
CC and a polymerase. The method is useful for detecting the cytosine  
CC methylation state and/or single nucleotide polymorphisms in genomic DNA,  
CC particularly for characterizing, classifying and/or differentiating renal  
CC and prostate cancers. The oligomers are useful as primer oligonucleotides  
CC for the amplification of any of the 112 DNA sequences of the invention.  
CC The set of oligomer probes is useful for detecting the cytosine  
CC methylation state and/or single nucleotide polymorphisms in any of the  
CC 112 chemically pretreated genomic DNA sequences. The method is also  
CC useful for identifying the tissue of origin of cancer cells. The method  
CC allows the classification, differentiation and/or diagnosis of cancer  
CC tissues using minute samples which would be inadequate for histological  
CC or cytological analysis. The present sequence represents one of the 112  
CC DNA sequences of the invention.  
XX  
SQ Sequence 7040 BP; 2018 A; 184 C; 1566 G; 3272 T; 0 U; 0 Other;  
Query Match 7.9%; Score 40.6; DB 8; Length 7040;  
Best Local Similarity 50.8%; Pred. No. 1.8; Mismatches 0; Gaps 0;  
Matches 97; Conservative 0; Indels 94; Indels 0; Gaps 0;  
QY 101 TACGTTTCTCCAGTTGTATTAAGACCTCCACAGATATAAGTCTATGCAACAAAGA 160  
DB 3586 TATATTTTAAAAAATCAATATTAATAAATCCCTTCTCTAATAACCTATATATTAATT 3527  
QY 161 AAATGTCATACATTCTCTTACTGCTCATTTATTTTTCATTAGATAGCCGGTTTTTACT 220  
DB 3526 ATATATAACCACTTTTATATATAACAATTCATTTTAAAAATATATTAACCTTTAAAAA 3467  
QY 221 ACAACTCAAAATTAAGATGAACGAATGATGGTTAGTACGTTTATTAAGAAGGTAT 280  
DB 3466 ATTAATTTAATTAACATTAACCAATTACAAAAACAATTTAAATCTATATAAACCTTAAAT 3407  
QY 281 AAAGTACTAT 291  
DB 3406 AAAATTTCTAT 3396  
RESULT 15  
ID ADE84191/c  
XX

ID ADE84191 standard; DNA; 7040 BP.  
 AC ADE84191;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Human lymphoid cell proliferative disorder gene derived DNA #127.  
 XX  
 KW ds: lymphoid cell proliferative disorder; methylation;  
 KW methylated CpG dinucleotide; single nucleotide polymorphism; SNP;  
 KW diffuse large B-cell lymphoma; mantle cell lymphoma;  
 KW chronic lymphocytic leukemia; small lymphocytic lymphoma;  
 KW follicular lymphoma; diagnosis; prognosis.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO2003044226-A2.  
 PD 30-MAY-2003.  
 XX  
 PF 25-NOV-2002; 2002WO-EP013265.  
 XX  
 PR 23-NOV-2001; 2001DE-01057491.  
 PR 28-DEC-2001; 2001DE-01064501.  
 XX  
 PA (EPIG-) EPIGENOMICS AG.  
 XX  
 PI Burger M, Caldwell C, Gene B, Becker E, Maier S, Nimnich I;  
 XX  
 DR WPI; 2003-457621/43.  
 XX  
 PT Detecting and differentiating between lymphoid cell proliferative  
 PT disorders comprises contacting a target nucleic acid with at least one  
 PT reagent that distinguishes between methylated and non-methylated CpG  
 PT dinucleotides.  
 XX  
 PS Claim 26: SEQ ID NO 187; 448bp; English.  
 XX  
 CC The invention relates to a method of detecting and differentiating  
 CC between lymphoid cell proliferative disorders associated with at least  
 CC one gene and/or their regulatory regions in a subject by contacting a  
 CC target nucleic acid in a biological sample obtained from the subject with  
 CC at least one reagent or series of reagents that distinguish between  
 CC methylated and non-methylated CpG dinucleotides within the target nucleic  
 CC acid. The genes and/or their regulatory regions are preferably selected  
 CC from MDRI, CSNR2B, ESR4, AR, CDK4, RB2, CDC25A, GRB beta, MYOD1, CDH3,  
 CC MYC1, ELK1, ABL1, APC, BCL2, CDH1, CDKN1A, CDKN1B, CDKN2A, CDKN2B, PDS,  
 CC GSTP1, HIC-1, MGMT, MHL1, MOS, MYC, PTEN, RBL2, TGFBR2, TP73, CDKN1C,  
 CC GSK3beta, ESR1, APAF1, BAK1, BAX or HOXA5. Oligomers, peptide nucleic  
 CC acid (PNA)-oligomers and/or isolated nucleic acids based on the sequences  
 CC of the genes are useful for detecting the methylation state of all the  
 CC CpG dinucleotides within one or more the sequences, or their complements,  
 CC for determining the cytosine methylation state and or single nucleotide  
 CC polymorphisms (SNPs), and for differentiating at least two of the medical  
 CC conditions such as diffuse large B-cell lymphoma, mantle cell lymphoma,  
 CC chronic lymphocytic leukemia, small lymphocytic lymphoma and follicular  
 CC lymphoma. They are also useful for detecting of a predisposition to,  
 CC differentiation of lymphoid cell proliferative disorder, diagnosis, prognosis, treating and/or  
 CC monitoring of lymphoid cell proliferative disorder. This sequence  
 CC represents a nucleic acid of a pretreated genomic DNA derived from the  
 CC above mentioned genes.  
 XX  
 SQ Sequence 7040 BP; 2018 A; 0 C; 1566 G; 3456 T; 0 U; 0 Other;

Query Match 7.9%; Score 40.6; DB 10; Length 7040;  
 Best Local Similarity 50.8%; Pred. No. 1.8;  
 Matches 97; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 101 TAGCTTTCCAGTTGATATAGACCTCCACAGTATAGTCTTATGCAACAGA 160  
 DB 3586 TATATTTTAAATAATCAATATATATATATATATATATATATATAT 3527  
 QY 161 AAATGCAATACATTCCTAGTCTCATATATATATATATATATATATAT 220

DB 3526 ATATATACACATTTTATATATATATATATATATATATATATATATAT 3467  
 QY 221 ACACTCAATTAAGATGACAGATGATGAGTGTAGTCTGTTTATAGAGAGTAAT 280  
 DB 3466 ATAAATTAATTAACATTAACATTTTACAAAACATTTTAAATCTATTAACCTTAAAT 3407  
 QY 281 AAGATACTAT 291  
 DB 3406 AAATTTCTAT 3396

Search completed: May 27, 2006, 19:15:57  
 Job time : 346.339 secs



funded by the NIH Centers of Excellence in Genomic Science (CEGS) initiative (<http://cegs.stanford.edu>). The clone was isolated from the BAC library CHORI-213 built by Pletier deJong in collaboration with the Stanford Genome Evolution Center (<http://www.chori.org/bacpac/>). Clones may be purchased from BACPAC Resources ([http://www.chori.org/bacpac/ordering\\_information.htm](http://www.chori.org/bacpac/ordering_information.htm)).

## ORIGIN

```
Query Match      9.3%; Score 47.6; DB 13; Length 1206;
Best Local Similarity 45.5%; Pred. No. 0.23;
Matches 170; Conservative 0; Mismatches 204; Indels 0; Gaps 0;

QY 11 TTCACTCATGCTATTATGAGTGAATAGATTCGACCAATTAAGATGACAAAG 70
Db 999 TTTAACTATATATTTTATGAAATACATATCTAATTAATTAAGAAATTAAGT 940
QY 71 CAGGATTTGGTTTATCTGTTGGTAAATACGTTTCTCCAGTTGATTAAGACCTC 130
Db 939 CTATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 880
QY 131 CCACAGATTAATAGTCCATGCAACAAGAAATGCAATCAATCTCTTATGCTCATTA 190
Db 879 TAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 820
QY 191 TTATTTTCAATTAAGTACCGGTTTCTTACCAACTCAATTAAGATGAACAGATGAATG 250
Db 819 ATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 760
QY 251 GGTAGTACGCTTTTAAAGAGATTAATTAATTAATTAATTAATTAATTAATTA 310
Db 759 TATATGAATTTCCATTAACAAATTAATTAATTAATTAATTAATTAATTAATTA 700
QY 311 GAGGAGAGATTCAGCAACAGTGTCTTACCAAGTGAAACCAATTAATTAAGTGA 370
Db 699 CAATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 640
QY 371 CCCCCCTCTTGAC 384
Db 639 CACACTTCATTAC 626
```

RESULT 2  
CNS014PO 987 bp DNA linear GSS 26-JUN-1999  
LOCUS Drosophila melanogaster genome survey sequence SP6 end of BAC  
DEFINITION BACN12P22 of DrosBAC library from Drosophila melanogaster (fruit  
fly), genomic survey sequence.  
ACCESSION AL104456.1 GI:5616067  
VERSION GSS.  
KEYWORDS Drosophila melanogaster (fruit fly)  
SOURCE Drosophila melanogaster  
ORGANISM Drosophila melanogaster  
Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 987)  
Genoscope.  
REFERENCE Direct Submission  
AUTHORS Submitted (23-JUN-1999) Genoscope - Centre National de Sequencage :  
TITLE BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)  
JOURNAL - Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the European Drosophila Genome Project (EDGP) -  
<http://www.edgp.ebi.ac.uk>. This Drosophila melanogaster BAC  
library (Dros BAC) was made by Alain Billaud at CEPH (Centre  
d'Etude du Polymorphisme Humain) with funding provided by a MC  
project grant. The DNA was prepared from embryos by Alain Bucheton  
and Genevieve Payan. It has been constructed in the vector  
pBelobAC11.  
location/Qualifiers  
1..987  
/organism="Drosophila melanogaster"

FEATURES  
source

/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACN12P22"  
/clone\_1ib="DrosBAC"  
/plasmid="pBelobAC11"  
/note="end : SP6"

## ORIGIN

```
Query Match      9.0%; Score 46; DB 14; Length 987;
Best Local Similarity 33.5%; Pred. No. 0.59;
Matches 78; Conservative 51; Mismatches 104; Indels 0; Gaps 0;

QY 136 AGTATTAAGTCTATGCAACAAGAAATGCAATATCTCTTACTCATATTAAT 195
Db 723 AAAAAAAMAAAMAAATTTTAAATTTTAAAGAAATTAATTAATTAATTTTAA 782
QY 196 TTCAATTAATGCGGTTTCTTACCACTCAATTAAGTGAACAGATGAATGGTTA 255
Db 783 TTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 842
QY 256 GTGACTGTTTAAAGAGATTAATTAAGATATCATCATCTTGAAGCAATGAAGAG 315
Db 843 TTTTATTAATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTA 902
QY 316 GAGGATTCAGCAACAGTGTCTTACCAAGTGAACCAATTAATTAAGT 368
Db 903 RAAMWMMWMTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 955
```

RESULT 3  
CNS000E/c 1101 bp DNA linear GSS 04-JUN-1999  
LOCUS Drosophila melanogaster genome survey sequence TE73 end of BAC #  
DEFINITION BACR28A22 of RPCI-98 library from Drosophila melanogaster (fruit  
fly), genomic survey sequence.  
ACCESSION AL067626.1 GI:4948726  
VERSION GSS.  
KEYWORDS Drosophila melanogaster (fruit fly)  
SOURCE Drosophila melanogaster  
ORGANISM Drosophila melanogaster  
Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 1101)  
Genoscope.  
REFERENCE Direct Submission  
AUTHORS Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
TITLE BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)  
JOURNAL - Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the Berkeley Drosophila Genome Project (BDGP).  
The BDGP is constructing a physical map of the Drosophila  
melanogaster genome using these BACs. For further information  
please see <http://www.fruitfly.org> The BDGP Drosophila  
melanogaster BAC library was prepared by Kazuo Oosawa and  
Aaron Mammeter in Pletier de Jong's laboratory in the Department of  
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,  
NY. The library is named RPCI-98 and was constructed by partial  
EcoRI digestion of Drosophila DNA provided by the BDGP from the  
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's  
P1 and EST libraries. A more detailed description of the library  
and how to order individual BAC clones, the entire library, or  
filters for hybridization from the BACPAC Resource Center can be  
found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).  
location/Qualifiers  
1..1101  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACR28A22"  
/clone\_1ib="RPCI-98"  
/note="end : TE73"

FEATURES  
source  
ORIGIN





Db	978	RDGGRKDKDKDRKXGDDDDKKGGKKKKAAKAAKATATXWDDWDDXDKXWMDGAKXDRKADD	919
Qy	228	AACGAAATGGAATGGTTAGTGAAGTCTTTATATAAGAAAGTAAATGAATCTATCATCAT	297
Db	918	DDGAGDKDDDEGKXADDDTDGTCKDDDDKDKWDDWDKAKGTWGDATWMAATDWWWWGWA	859
Qy	298	TTGAGGCATATTAAGGAGAGAGATTCAGCAAAAGCTGGCTTACAAATGGGAAAAACAGT	357
Db	858	DADWTTWDAADWDAADDRWDAADWMAKWDAMAKGRTADRRDWDGDRAGKRGARRRDRK	799
Qy	358	TAACTAAAGTGA	370
Db	798	RADDKRDAAADRD	786
RESULT 8			
LOCUS	CL108838/c	1158 bp	DNA linear GSS 05-JAN-2004
DEFINITION	ISB1-5012_Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-5012,		
ACCESSION	CL108838		
VERSION	CL108838.1		
KEYWORDS	GI:40602473		
SOURCE	GSS.		
ORGANISM	Xenopus tropicalis (western clawed frog)		
REFERENCE	Xenopus tropicalis		
AUTHORS	Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus; Silurana.		
TITLE	Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T., Mendis, E. and Wilson, R.		
JOURNAL	A physical map of the xenopus tropicalis genome		
COMMENT	Unpublished (2003)		
FEATURES			
source	Contact: Richard K Wilson Genome Sequencing Center Washington University School of Medicine Email: submision@watson.wustl.edu Insert Length: 7500 Std Error: 0.00 Seq primer: Sp6 ATTTAGGTGACACTATAG Class: BAC ends High quality sequence start: 36 High quality sequence stop: 766. Location/Qualifiers		
ORIGIN	1..1158		
	/organism="Xenopus tropicalis"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:8364"		
	/clone="ISB1-5012"		
	/clone_11b="ISB1"		
	/note="Vector: pBelobAC11; ISB-1 Xenopus tropicalis BAC Library Segment 1"		
Query Match	8.6%	Score 43.8;	DB 12; Length 1158;
Best Local Similarity	52.5%;	Pred. No. 2.3;	
Matches	96; Conservative	0; Mismatches	87; Indels 0; Gaps 0;
Qy	115	TTGTATTAAGACCCCTCCACGAGTAAAGTCATGACAAAGAAATGTCATACAT	174
Db	781	TGGAATAGGATCTCACAGAAATATGTAGGCTTAAAGAAAGAAAAAGAAAAA	722
Qy	175	TCTCTAGTCTCATATATATTTTCATTAGATAGCGGTTTTTTTACTACACTCAATAG	234
Db	721	GCTGTTGGTAAATTTGTTAATGTTTATAGGCGCTGTTAAGACTCAAGTTAATGAA	662
Qy	235	ATGAACAGAAATGATGGGTTAGTGACTGTTTATATAAGAGCTAATAAGTACTATCAT	294
Db	661	GTTAAATACTAAAGTTATGCACTATATCCATAGTCAATATCTTATTAATTAACCT	602
Qy	295	CAT	297
Db	601	TAT	599

LOCUS	CG769354	927 bp	DNA	linear	GSS 29-OCT-2003
DEFINITION	CG769354	TcB41.1.H04.SP6	Tribolium BAC library	Tribolium castaneum genomic,	
ACCESSION	CG769354				
VERSION	CG769354.1	GI:38022533			
KEYWORDS	GSS.				
SOURCE	Tribolium castaneum (red flour beetle)				
ORGANISM	Tribolium castaneum				
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformata; Tenebrionidae; Tribolium.				
AUTHORS	1 (bases 1 to 927)				
TITLE	Savard,J. and Tautz,D.				
JOURNAL	Tribolium castaneum BAC-ends sequencing project				
COMMENT	Unpublished (2003)				
FEATURES	Contact: Savard, J.				
source	Abteilung fur Evolutionsgenetik, AG Tautz				
	Institut fur Genetik, Universitat zu Koln				
	Weyertal 121, 50931 Koln, Germany				
	Tel: 49 221 470 6911				
	Fax: 49 221 470 5975				
	Email: savard@uni-koeln.de				
	Class: BAC ends.				
	Location/Qualifiers				
	1..927				
	/organism="Tribolium castaneum"				
	/mol_type="genomic DNA"				
	/strain="GA-2"				
	/db_xref="taxon:7070"				
	/clone_lib="Tribolium BAC library"				
	/note="Vector: PBAC3.6; Site 1: EcoRI; Site 2: EcoRI;				
	library constructed by Exelixis Inc."				
ORIGIN					
	Query Match	8.5%	Score 43.6;	DB 12;	Length 927;
	Best Local Similarity	53.8%	Pred. No. 2.5;		
	Matches 112;	Conservative 0;	Mismatches 94;	Indels 2;	Gaps 1;
QY	138 TATAAATCCATGCAACAAGAAATGCAATACATTCCTTAGTCCATTTATTTT	197			
DB	26 TATTAATCGATGTTAATTTAAACGTCATTAATTTGATTTGTTACATGCAACTA	85			
QY	198 CATTAGATPACCGGTTTTTTACTACACTCAATTAAGATGAACAGATGAGGTAGT	257			
DB	86 AATTGCAATTTACACTGACGACAGAAATACCTTTATTTTAAAGAAAGAACGATTT	145			
QY	258 GACTGTTATTAAGAGAGATTAATAA--GATACTATCATCATTTGAGGCAATAGGAGG	315			
DB	146 AAACGGCAATTAACAGATTAATTAACCCGATTCATGTTATTTCAAGTCAACGTTGAGG	205			
QY	316 GAGAGATTCAGCAACAGGTGCTTACA	343			
DB	206 AAGAGATTACAGAAACCGCCTGTTTCA	233			
RESULT 10					
LOCUS	AV008603	443 bp	mRNA	linear	EST 25-AUG-1999
DEFINITION	AV008603 Mus musculus 18-day embryo C57BL/6J Mus musculus cDNA				
ACCESSION	AV008603				
VERSION	AV008603.1	GI:4785590			
KEYWORDS	EST.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;				
	Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 443)				

**AUTHORS** Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K., Akahira, S., Akiyama, J., Fukuda, S., Fukushima, Y., Funayama, T., Hara, A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M., Kawai, J., Kikuchi, N., Kojima, Y., Matsuyama, T., Nitsuna, H., Oda, H., Owa, C., Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tateo, M., Tomaru, Y., Tomimaga, N., Watanabe, S., Yagame, M., Yamamura, T., Yokota, T., Yoshino, M., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

**TITLE** RIKEN Mouse ESTs

**JOURNAL** Unpublished (1999)

**COMMENT** Contact: Chie Owa  
Genome Science Laboratory  
RIKEN  
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel.: 81-298-36-9145  
Fax: 81-298-36-9098  
Email: genome-resortc.riken.go.jp  
Thermolabile and thermostabilization of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))  
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))  
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

**FEATURES**  
source  
1..443  
/organism="Mus musculus"  
/mol\_type="rRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="1110013A05"  
/sex="mixed"  
/dev\_stage="18-day embryo"  
/clone\_lib="Mus musculus 18-day embryo C57BL/6J"

**ORIGIN**

Query Match 8.4%; Score 43.2; DB 1; Length 443;  
Best Local Similarity 55.3%; Pred. No. 2.8;  
Matches 84; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

154 ACAAGAAAGTCAATCATCTCTAGTCATATTTATTTATAGATGAGCCGTT 213  
3 ACTAAATTTCTCCAAACATCCCTTAATTTTCATTTAGATGCGTGAAGAA 62  
214 TTTTACTCAACTCAATAGATGAGATGAGATGAGTGGTGTAGTCTTTATTAAGA 273  
63 ATGAACCTTCGAGAGAACAGAACATGAAAGTGGTTATATGACTGTTTGAATTA 122  
274 GAGTAAATAAGTACTATCATCATTTGAGGCA 305  
123 AAGATAGGTGCTCATAGAGTTTATGATTC 154

**RESULT 11**

LOCUS DX034017 784 bp DNA linear GSS 10-JAN-2006  
DEFINITION KBAB031L03F KBrB, Brassica rapa BamHI BAC library Brassica rapa subsp. pekinensis genomic clone KBAB031L03, genomic survey sequence.  
ACCESSION DX034017  
VERSION DX034017  
KEYWORDS DX034017.1 GI:84728314  
ORGANISM GSS.  
Brassica rapa subsp. pekinensis  
Brassica rapa subsp. pekinensis  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
REFERENCE 1 (bases 1 to 784)  
Yang, T.J., Kwon, S.J., Kim, J.A., Kim, J.S., Lim, K.B., Jin, M., Park, J.Y., Lim, M.H., Kim, H.I., Choi, B.S., Seol, Y.J., Park, D.S., Hahn, J.H. and Park, B.S.  
End sequence of Brassica rapa BamHI (KBrB) BAC clone  
JOURNAL Unpublished (2005)

**COMMENT** Contact: Beom-Seok Park  
Beom-Seok Park  
National Institute of Agricultural Biotechnology  
225 Seodun-Dong, Suwon, 441-707, Korea  
Tel.: +82-31-299-1670  
Fax: +82-31-299-1672  
Email: pbeom@da.riken.go.jp  
BAC end sequence of Brassica rapa subsp. pekinensis BamHI BAC clone KBAB031L03

**TITLE** Seq primer: 77  
Class: BAC ends.

**FEATURES**  
source  
1..784  
/organism="Brassica rapa subsp. pekinensis"  
/mol\_type="genomic DNA"  
/cultivar="Chifu"  
/sub\_species="pekinensis"  
/db\_xref="taxon:51351"  
/clone="KBAB031L03"  
/lab\_host="E. coli DH10B"  
/clone\_lib="KBrB, Brassica rapa BamHI BAC library"  
/note="Vector: pCUGIBAC1; Site 1: BamHI; Brassica rapa subsp. pekinensis var. Chifu BAC library (KBrB BAC) is provided by Yong-Pyo Lim (CNU)."

**ORIGIN**

Query Match 8.4%; Score 43.2; DB 14; Length 784;  
Best Local Similarity 50.0%; Pred. No. 3.1;  
Matches 108; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

58 AGAATGACAAAGCAGAGATTGTTGTTTATCTGTGGTAAATACGTTTCTCCAGTTG 117  
368 AAAATCGCGCCAGACAGATTAGGTTTATTTTATTAATTAAGACATAGTTTACGGGAA 427  
118 TATTAAGACCTCCACAGATTAAGTCTATGCAACAAAGAAATGTCAATCATCTT 177  
428 CAAATATATATTAATTAATCAAAATTAATTAACGATTAATATATGTCACATTAATTT 487  
178 CTTAGTCTCATTTATTTTATTTATTTATTTATTTTATTTTATTTTATTTTATTTT 237  
488 TCACATGATTAATTTATTTTATTTATTTATTTATTTATTTATTTTATTTTATTTT 547  
238 AACGAAATGAATGGGTTAGTACTGTTTATTAAGAA 273  
548 AACATATTAATAATGTTTATTTTATTTTATTTTATTTATTTATTTATTTATTTATTT 583

**RESULT 12**

LOCUS DX077691 886 bp DNA linear GSS 10-JAN-2006  
DEFINITION KBAB089F21F KBrB, Brassica rapa BamHI BAC library Brassica rapa subsp. pekinensis genomic clone KBAB089F21, genomic survey sequence.  
ACCESSION DX077691  
VERSION DX077691.1 GI:84771987  
KEYWORDS DX077691  
ORGANISM GSS.  
Brassica rapa subsp. pekinensis  
Brassica rapa subsp. pekinensis  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
REFERENCE 1 (bases 1 to 886)  
Yang, T.J., Kwon, S.J., Kim, J.A., Kim, J.S., Lim, K.B., Jin, M., Park, J.Y., Lim, M.H., Kim, H.I., Choi, B.S., Seol, Y.J., Park, D.S., Hahn, J.H. and Park, B.S.  
End sequence of Brassica rapa BamHI (KBrB) BAC clone  
JOURNAL Unpublished (2005)  
Contact: Beom-Seok Park  
National Institute of Agricultural Biotechnology  
225 Seodun-Dong, Suwon, 441-707, Korea  
Tel.: +82-31-299-1670  
Fax: +82-31-299-1672





D <sub>b</sub>	234	A T A A T T G A C T T T T G T T T T G T T T T G A G T G T A T A T A C G T T G A A G T T A A T T T T T T T T A	293
O <sub>y</sub>	229	A A T A G A T G A C A G A T G A T G G G T A G T A G C A C T G T T A T A A A G A A G T A T A A A G A T A C	288
D <sub>b</sub>	294	T C A A A C A C A A T T G C C A G A A A T A G A T G T T G G A T C A T A T A T T T G C A A A A T A T A T A T A C	353
O <sub>y</sub>	289	T A T C A T C A T T T G A G G C A A T A A G G A G G A G A G A T T C A G C A A A C A G T G T G T T A C A A G T G	348
D <sub>b</sub>	334	T T G G T C A T T T A A A A A G T T A A B A C A C G T T A A T A T G T C A A A A A A A A T G T C A T G A A A T T G G	413
O <sub>y</sub>	349	A A A A C A A G T T A A C T A A A G T G A C C C C C T T C T T G A C A A G A T C A A T G C A C A G T T G	403
D <sub>b</sub>	414	C C C A T T A G T T A A A A A C A T A T A G T T G C T C A A T A C A A A A T T A A G C C C A C A T T T G	468

```

Db      303  TACAAAAGAAAGTAAATTCACCTGGAAATTAATGTTTTATGANGTAAGATATTCAG 244
QY      208  CCGGTTTTTACTACACTCAATATAGATGAACAGAAAGATGGGTAGAGACTGTTAT 267
Db      243  AAAGTAAATATCCACCAAGCTGTAGTAAATGAAGAAATGGATGTTTTGAGAAATGTTCA 184
QY      268  AAAGAGAGATTAATAAGATACTATCATCTTGTAGGCAATTAAGGAGAGAGATTACG 327
Db      183  CATTAATAAACATTAATATATACCATCATCAATTACATCATTTGGAGTTGATGTTTTATC 124
QY      328  AAACAGTGTCTTACAATGGAAGAAACCAATTAACTAA 366
Db      123  ATTGTGATGTGAAGAAATTTAAAGAAATATTAACA 85

Search completed: May 28, 2006, 14:46:43
Job time : 2628.13 secs

```

Search completed: May 28, 2006, 14:46:43  
Job time : 2628.13 secs

LOCUS	BH137539	879 bp	DNA	linear	GS5 07-AUG-2001
DEFINITION	ENTN1A83TR Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, genomic survey sequence.				
ACCESSION	BH137539				
VERSION	BH137539.1	GI:15096600			
KEYWORDS	GS5.				
SOURCE	Entamoeba histolytica				
ORGANISM	Entamoeba histolytica				
REFERENCE	Entamoeba histolytica				
AUTHORS	Eukaryota; Entamoebidae; Entamoeba.				
TITLE	1 (bases 1 to 879)				
JOURNAL	Loftus, B., Wang, Z., Van Aken, S. and Fraser, C.				
COMMENT	Determination of clone end sequences from Entamoeba histolytica				
	HM1:IMSS sheared DNA library (2001)				
	Unpublished (2001)				
	Contract: Brendan J Loftus				
	Department of Pathologic Microbiology, University of Maryland School of Medicine, Baltimore, Maryland, USA				

FEATURES	Location/Qualifiers
source	1. .879

/clone\_libs="Entamoeba histolytica sheared DNA"  
/note="Vector: pROS1; Site 1: Bst I; Constructed at The  
Institute for Genomic Research (TIGR), Rockville, MD.  
Genomic DNA isolated from broth cultures of *E. histolytica*  
using a method described by Clark and Diamond (Clark,  
C.G., and Diamond, L.S. (1993) *Entamoeba histolytica*: a  
method for isolate identification. Exp. Parasitol.  
77:450.). The DNA was mechanically sheared to give a  
tight size distribution (~2 kb). The v + i method used for  
the library construction is described in detail in Smith,  
H.O. and Venter, J.C. (Making small insert libraries for  
whole genome shotgun sequencing projects. In Genome  
Sequencing: A Practical Approach, eds. M. Vaubin and B.  
Sartell, Oxford University Press, 1999)."

## ORIGIN

Query Match	8.4%	Score 43;	DB 11;	Length 879;
Best Local Similarity	49.8%	Pred. No. 3.5;		
Matches 109;	Conservative	0;	Mismatches 110;	Indels 0;
				Gaps 0;

148 TATGCACAAAGAAATGTCATATACATTCCTTAGTCCTATATTATTTTCAATTAGATAG 207

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd

OM nucleic - nucleic search, using sw model

Run on: May 27, 2006, 18:52:01 ; Search time 102.294 Seconds  
(without alignments)  
9365.256 Million cell updates/sec

Title: US-10-099-663-1\_COPY\_1115\_1626  
Perfect score: 513

Sequence: 1 cagtggatcctcatctcat.....gtctgcctacagacagaag 512

scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 93555401 residues

Total number of hits satisfying chosen parameters: 2807332

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Database : Issued\_Patents\_NA:\*

```

1:  /EMC_Celerra_S1D33/ptodatea/2/ina/1_COMB.seg.*
2:  /EMC_Celerra_S1D53/ptodatea/2/ina/5_COMB.seg.*
3:  /EMC_Celerra_S1D53/ptodatea/2/ina/6A_COMB.seg.*
4:  /EMC_Celerra_S1D53/ptodatea/2/ina/6B_COMB.seg.*
5:  /EMC_Celerra_S1D53/ptodatea/2/ina/7_COMB.seg.*
6:  /EMC_Celerra_S1D53/ptodatea/2/ina/H_COMB.seg.*
7:  /EMC_Celerra_S1D53/ptodatea/2/ina/PP_COMB.seg.*
8:  /EMC_Celerra_S1D53/ptodatea/2/ina/PP_COMB.seg.*
9:  /EMC_Celerra_S1D53/ptodatea/2/ina/RE_COMB.seg.*
10: /EMC_Celerra_S1D53/ptodatea/2/ina/backfilled.seg.*

```

**Pred.** No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match			ID	Description
		Length	DB			
1	45.4	8.9	114.1	3	US-09-806-7088-22	Sequence 22, Appl
2	42.2	8.2	125.7	3	US-09-270-767-15058	Sequence 15058, A
3	40	7.8	73.3	3	US-09-270-767-10136	Sequence 10136, A
4	39.2	7.7	486.0	3	US-09-270-767-10504	Sequence 10504, A
5	38.4	7.5	323.92	3	US-09-662-2548-27	Sequence 27, Appl
6	38.4	7.5	5000.0	3	US-09-662-2548-23	Sequence 23, Appl
7	38.2	7.5	6573.7	3	US-09-782-1788-17	Sequence 17, Appl
8	38	7.4	46068.1	2	US-09-782-988-1	Sequence 1, Appl
9	37.8	7.4	72.8	2	US-08-232-463-14	Sequence 14, Appl
10	37.6	7.3	60.1	3	US-09-949-016-123489	Sequence 123489, A
11	37.6	7.3	4435.3	3	US-09-949-016-15302	Sequence 15302, A
12	37.6	7.3	10821.4	4	US-09-531-120-221	Sequence 211, App
13	37.4	7.3	66.3	3	US-08-998-416-187	Sequence 187, App
14	37.4	7.3	696	3	US-08-998-416-779	Sequence 779, App
15	37.4	7.3	719	3	US-08-998-416-1138	Sequence 1138, App
16	37.4	7.3	856	3	US-08-998-416-289	Sequence 289, App
17	37	7.2	832	3	US-09-621-976-2813	Sequence 2813, Ap
18	36.8	7.2	4311.7	3	US-09-949-016-17589	Sequence 17589, A
19	36.6	7.1	60.1	3	US-09-949-016-165407	Sequence 165407, A
20	36.6	7.1	601	3	US-09-949-016-1565464	Sequence 165464, A
21	36.6	7.1	2754.3	3	US-09-949-016-17455	Sequence 17455, A
22	36.6	7.1	3664.3	3	US-09-949-016-11860	Sequence 11860, A
23	36.6	7.1	3682.1	3	US-09-949-016-16403	Sequence 16403, A

C	24	36.6	7.1	36621	3	US-09-949-016-16404	A	Sequence 16404, A
C	25	36.2	7.1	601	3	US-09-949-016-17378	A	Sequence 17378, A
C	26	36.2	7.1	601	3	US-09-949-016-17375	A	Sequence 17375, A
C	27	36.2	7.1	832	3	US-09-622-976-2813	A	Sequence 2813, Ap
C	28	36.2	7.1	1587	3	US-09-710-279-1929	A	Sequence 1929, Ap
C	29	36.2	7.1	1878	3	US-09-134-001C-1667	Ap	Sequence 1667, Ap
C	30	36.2	7.1	3000	3	US-09-710-279-3907	Ap	Sequence 3907, Ap
C	31	36.2	7.1	3368	3	US-09-710-279-9459	Ap	Sequence 4169, Ap
C	32	36.2	7.1	81819	3	US-09-949-016-1661	Ap	Sequence 1661, A
C	33	36.2	7.1	81819	3	US-09-949-016-1662	A	Sequence 1662, A
C	34	36	7.0	1055	3	US-09-806-708B-23	A	Sequence 23, Ap1
C	35	36	7.0	168394	3	US-09-949-016-13002	A	Sequence 13002, A
C	36	36	7.0	786431	3	US-09-751-389-3	Ap1	Sequence 3, Ap1
C	37	35.8	7.0	714	3	US-09-662-254B-64	Ap1	Sequence 64, Ap1
C	38	35.8	7.0	1557	3	US-09-270-767-28519	Ap1	Sequence 28519, A
C	39	35.8	7.0	3128	3	US-09-270-767-12706	Ap1	Sequence 12706, A
C	40	35.8	7.0	50000	3	US-09-662-254B-26	Ap1	Sequence 26, Ap1
C	41	35.8	7.0	193956	3	US-09-949-016-14382	A	Sequence 14382, A
C	42	35.8	7.0	1230025	3	US-09-198-452A-1	Ap1	Sequence 1, Ap1
C	43	35.8	7.0	1230230	3	US-09-438-185A-1	Ap1	Sequence 1, Ap1
C	44	35.6	7.0	601	3	US-09-949-016-44395	A	Sequence 44395, A
C	45	35.6	7.0	601	3	US-09-949-016-58898	A	Sequence 58898, A

## ALIGNMENTS

```

RESULT 1
US-09-806-708B-22
Sequence 22; Application US/09806708B
Patent No. 6784342
GENERAL INFORMATION:
APPLICANT: The University of British Columbia
TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
FILE REFERENCE: 4810-58741
CURRENT APPLICATION NUMBER: US/09/806,708B
CURRENT FILING DATE: 2001-04-03
PRIORITY APPLICATION NUMBER: US 60/147,133
PRIOR FILING DATE: 1999-08-04
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.0
SEQ ID NO 22
LENGTH: 1141
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
NAME/KEY: promoter
LOCATION: (1)..(1141)
OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. PAB1 promoters
US-09-806-708B-22

Query Match      8.9%; Score 45.4; DB 3; Length 1141;
Best Local Similarity 9.9%; Pred. No. 0.0039;
Matches 46; Conservative 197; Mismatches 221; Indels 0; Gaps 0;

QY    12 TCATCTCATGCTTATTATGTGAGTAGAATTCACCAATTTAGAATGACAAGC   71
       |||::||::||::||::||::||::||::||::||::||::||::||::||
DB    110 YEMTMNNKKGKIGMRFRRYRWRMBDVTVDHHVYTANNNAWTMMCMWDXDKETRMWWKKN   169

QY    72 AGGAGTTTGCGTTTATCTGTTGGGTAATAAGCTTTTCCAGTGTATAAAGACCCTCC   131
       ||::||::||::||::||::||::||::||::||::||::||::||::||
DB    170 NNAATGWDDDKYYMNNNNNGCBVTWVRKYKTRDWBSBKMNNGMBWVKWSYDVYYWVVAV   229

QY    132 CACGAGTAAAGTCCATGCAACAAGAAAATGTCATACATTCCTCTAGTCTCATTTAT   191
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB    230 WDMCKRKRRRWRTGRGRNRNVVAWBTHHRRRYNNGTBMAAYRRWTANNNNNAKAMC     289

QY    192 TATTTTCATTAGATAGCCGGGTTTTTACTACAACTCAAATTAAGATGAACAGATGAATGG   251
       ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB    290 KRAKYGMNRABVNSCTTWKSRTKVRTSCWANNCRADDAKHDKMKMSAAMGVYNNN     349

QY    252 GTTAGGCATGTTATTAAGAGAGTAATTAAGATCTCATCATTTGAGCAATTAAG     311
       :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```
Db 350 NNNNNNTYKKBHBAWMDVWHSAMKMHANAHSRKKTBYKKTWVNNNGTTMKR 409
Qy 312 GAGGAGAGATTGAGCAAAAGCTGTGCTTACAGTGAAGAAAACAAGTTAACTAAAGTGAC 371
Db 410 MAAWYWKMDMBWYKNNNNNGRTYYGTYKTKKWTYYKWKVANNCKRAMDHKTCTHNN 469
Qy 372 CCCCCCTCCCTCAAGATCATGCGACAGTTGAGCTTGAAGCAGACGACATCATATGTA 431
Db 470 TTTWIKTKYNNNCYKSKNTNGSHBBAAVTYTWYMMRRYAHANNNNNDYWMRACCTYK 529
Qy 422 ATTGCTTTCCTGATAAGCTGCTTCATTAATTCCTTTGCAAGC 475
Db 530 YBVCCKWNNNYAAWYTKSSWNTSYRYMKTNNSWRMSDTRSM 573
```

## RESULT 2

```
US-09-270-767-15058/C
; Sequence 15058, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 15058
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-15058
```

```
Query Match 8.2%; Score 42.2; DB 3; Length 1257;
Best Local Similarity 53.3%; Pred. No. 0.033;
Matches 89; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
```

```
Qy 73 GAAATTTGCTTTTATCTGTGGTAATAAGTTTTCTCAGTTGATTAAGACCTCCC 132
Db 364 GAGTTTATGATTTAATTTAATTTCACTCAACACAGACACGACCTACTC 305
Qy 133 ACCAGTATAAGTCTATGCAACAAGAAAGTCAATCTCTTGTCTCATTTT 192
Db 304 ACAGCAAAAACGTACAAGAGAGAAAGAAATTAAGAGTGATTTCTTCAATA 245
Qy 193 ATTTTATGATATGACCGGTTTTTATCTAACAATCAATTAAGATGA 239
Db 244 TGTATTGCAATTAAGGTGTGGCATTCATATCAATATAAATAAGTA 198
```

## RESULT 3

```
US-09-270-767-10136/C
; Sequence 10136, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 10136
; LENGTH: 733
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-10136
```

```
Query Match 7.8%; Score 40; DB 3; Length 733;
Best Local Similarity 50.5%; Pred. No. 0.12;
Matches 97; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
```

```
Qy 48 TCACCAATTGAAATGACAAAGAGATTGTGTTTATCTGTGGTAATAAGCTT 107
```

```
Db 306 TAAATATAATGCAAGTTATATTTGAGTTTATATTTATTTTAAGTATTTCACT 247
Qy 108 TCTCCAGTTGTATTAAGACCTCCACAGTATAAGTCTATGCAACAAAGAAATGTC 167
Db 246 GCAACGACGACGACGACCTACTACGACAAAGACGACAAAGAGAAAGGATA 187
Qy 168 AATACATCTCTTAATCTCATATATTTTGAATAGATAGAGCGGTTTTTACTCACTC 227
Db 186 AAAAAAGTGATTTCTCTTCAATATGTTTATGCAATAAAGGTGGCCATTCATATC 127
Qy 228 AATATAGTGA 239
Db 126 AATATATAAGTA 115
```

## RESULT 4

```
US-09-270-767-10504/C
; Sequence 10504, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 10504
; LENGTH: 4860
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-10504
```

```
Query Match 7.7%; Score 39.2; DB 3; Length 4860;
Best Local Similarity 46.6%; Pred. No. 0.39;
Matches 125; Conservative 0; Mismatches 143; Indels 0; Gaps 0;
```

```
Qy 97 TAAATAGTTTTCTCCAGTTGTATAAGACCTCCACAGTATAAGTCTATGCAACA 156
Db 1390 TAAATCGATTTATTTTGTATTAATATCAACGATTAATTAATGGAAGCATTAAT 1331
Qy 157 AAGAAATGCAATACATCTCTAGTCTCATTTATTTTATAGTATGAGCGGTTTTT 216
Db 1330 AATTAATTTATTTTATTAATATGCGCATATGCTTAATAATTTTTCGCTTACA 1271
Qy 217 TACTACACTCAATTAAGATGAACAGATGAATGGTTAGTACTGTTATTAAGAGAG 276
Db 1270 TTTTGTACGAATTAATTTGTAATTTACATTTGATTTGATTTTATTAAGATGC 1211
Qy 277 TAATAAGATACATCATATTTGAGCAATTAAGGAGAGAGATTCAGCAACAGTGT 336
Db 1210 CAAATAGTTTATCTTCAATATGAGGTAAGAAAGATTAATTAACCAATTAAGTTC 1151
Qy 337 GCTTACAGTGAACAAAGTTAACTA 364
Db 1150 ATTTTATGTTAAATTAAGTACATGCGA 1123
```

## RESULT 5

```
US-09-662-254B-27
; Sequence 27, Application US/09662254B
; Patent No. 6933145
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Li, Yi
; APPLICANT: Bawden, Alison Louise
; TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous
; FILE REFERENCE: UF-221C1X01
; CURRENT APPLICATION NUMBER: US/09/662,254B
; CURRENT FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 09/086,651
```

;; PRIOR FILING DATE: 1998-05-29  
;; PRIOR APPLICATION NUMBER: 60/224,479  
;; PRIOR FILING DATE: 2000-08-10  
;; NUMBER OF SEQ ID NOS: 80  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 27  
;; LENGTH: 32392  
;; TYPE: DNA  
;; ORGANISM: Ambacta moorei entomopoxvirus  
US-09-662-254B-27

Query Match 7.5%; Score 38.4; DB 3; Length 32392;  
Best Local Similarity 46.3%; Pred. No. 1.3;  
Matches 126; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 19 ATGCTATTATGAGTGAATGATTCTCACCACCAATTGAAATGACACAGAGATT 78  
DB 23294 ATGGTTATGATGTTGATTTTCTACTAAAGCGGAATTATATATGAAAAAGTGACT 23353  
QY 79 TGTGTTTATCTGTTGGTAATACGTTTCTCCAGTTGATTAAGACCCCTCCACAGT 138  
DB 23354 GGTGTCCTCTATGCTTATATATATATATATATATATATATATATATATAT 23413  
QY 139 ATAAAGTCTATGCAACAAAGAAATGCAATACATCTCTAGTCTATATATTTTC 198  
DB 23414 AAAAGGAACAAATGATGAAATATACATTAATATGATACGAATATATTTT 23473  
QY 199 ATTAATAGCCGGTTTTTTTACTACACTCAATTAAGATGAACAGATGAATGGTTAGT 258  
DB 23474 ATTAATACAAAGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 23533  
QY 259 ACTGTTTATTAAGAGATATATAGATCTACTA 290  
DB 23534 ATGCTTTAAATAATATATTAACAAATATATA 23565

RESULT 6  
US-09-662-254B-23/C  
; Sequence 23, Application US/09662254B  
; Patent No. 6933145  
; GENERAL INFORMATION:  
; APPLICANT: Moyer, Richard W.  
; APPLICANT: Li, Yi  
; APPLICANT: Bawden, Alison Louise  
; TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous  
; FILE REFERENCE: UF-221C1X1  
; CURRENT APPLICATION NUMBER: US/09/662,254B  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 09/086,651  
; PRIOR FILING DATE: 1998-05-29  
; PRIOR APPLICATION NUMBER: 60/224,479  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 23  
; LENGTH: 50000  
; TYPE: DNA  
; ORGANISM: Ambacta moorei entomopoxvirus  
US-09-662-254B-23

Query Match 7.5%; Score 38.4; DB 3; Length 50000;  
Best Local Similarity 46.3%; Pred. No. 1.5; Indels 0; Gaps 0;  
Matches 126; Conservative 0; Mismatches 146; Indels 0; Gaps 0;  
QY 19 ATGCTATTATGAGTGAATGATTCTCACCACCAATTGAAATGACACAGAGATT 78  
DB 9099 ATGGTTATGATGTTGATTTTCTACTAAAGCGGAATTATATATGAAAAAGTGACT 9040  
QY 79 TGTGTTTATCTGTTGGTAATACGTTTCTCCAGTTGATTAAGACCCCTCCACAGT 138  
DB 9039 GGTGTCCTCTATGCTTATATATATATATATATATATATATATATATATAT 8980

QY 139 ATAAAGTCTATGCAACAAAGAAATGCAATACATCTCTAGTCTATATATTTTC 198  
DB 8979 AAAAGGAACAAATGATGAAATATATATATATATATATATATATATATATAT 8920  
QY 199 ATTAATAGCCGGTTTTTTTACTACACTCAATTAAGATGAACAGATGAATGGTTAGT 258  
DB 8919 ATTAATACAAAGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8860  
QY 259 ACTGTTTATTAAGAGATATATAGATCTACTA 290  
DB 8859 ATGCTTTAAATAATATATTAACAAATATATA 8828

RESULT 7  
US-09-782-378A-17/C  
; Sequence 17, Application US/09782378A  
; Patent No. 6916635  
; GENERAL INFORMATION:  
; APPLICANT: Hearing, Patrick  
; APPLICANT: Bahou, Nadie  
; APPLICANT: Sandalon, Ziv  
; APPLICANT: Gnatelko, Dmitri  
; TITLE OF INVENTION: Adenoviral Vectors  
; FILE REFERENCE: STONYB-04370  
; CURRENT APPLICATION NUMBER: US/09/782,378A  
; PRIOR FILING DATE: 2001-02-12  
; PRIOR APPLICATION NUMBER: 60/237,747  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 17  
; LENGTH: 56737  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-782-378A-17

Query Match 7.5%; Score 38.2; DB 3; Length 56737;  
Best Local Similarity 52.9%; Pred. No. 1.8;  
Matches 82; Conservative 0; Mismatches 73; Indels 0; Gaps 0;  
QY 156 AAAAGAAATGCAATACATCTCTAGTCTATATATATTTTCAATGACCGGTTTT 215  
DB 18242 AAATAAATATGCAAGCAATTTCTACCACTGAAATGATCTTATATACATGAATGCTTT 18183  
QY 216 TTACTACAACTCAATTAAGATGAACAGATGAATGGTTAGTGAAGTATTAAGAAGA 275  
DB 18182 TTAAGAAATATTAAGTTAAAGTTAACTCCCTATTTTGTCTGTTTGTCTTATCTAATAA 18123  
QY 276 GTAATTAAGATCTATCATCATTTGAGGCAATTAAG 310  
DB 18122 ACATTTCGACAAATCCCAAGATTAAGTATTAAG 18088

RESULT 8  
US-09-790-988-1  
; Sequence 1, Application US/09790988  
; Patent No. 6632935  
; GENERAL INFORMATION:  
; APPLICANT: SHIGENOBU, SHUJI  
; APPLICANT: WATANABE, HIDEKI  
; APPLICANT: HATTORI, MASAHIRA  
; APPLICANT: SAKAKI, YOSHIYUKI  
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS  
; FILE REFERENCE: 081356/0159  
; CURRENT APPLICATION NUMBER: US/09/790,988  
; PRIOR FILING DATE: 2001-02-23  
; PRIOR APPLICATION NUMBER: JP2000-1071160  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 640681  
; TYPE: DNA

ORGANISM: Buchnera sp.  
US-09-790-988-1

Query Match 7.4%; Score 38; DB 3; Length 640681;  
Best Local Similarity 46.6%; Pred. No. 5;  
Matches 122; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

24 TATTATGTGAGTAAATGATTTCTCACCATTGATGACAAAGAGAGATTGTGT 83  
189923 TATTATTTCTGAGAAAGAAATTAATGATTTCTCACCATTGATGACAAAGAGATTGTGT 189982  
QY 84 TTTATCTTGGGTAATAACGTTTCTCCAGTTGTAATAAGACCCCTCCACAGTATATA 143  
Db 189983 TATGCAAAATTAGAAATTTCTCCAGTATATATAATAAGAAAAAAATTA 190042  
QY 144 GTCCATGCAACAAAGAAATGCAATATCATCTCTAGTCTCATTTATTTTCATTAG 203  
Db 190043 TTTTCTTTTAAACAAATAATATAATAATATTTAGTGTATTTTATCATCAATTA 190102  
QY 204 ATACCCGTTTCTTACTACACTCAATTAAGATGACAAATGAGTGGTTAGTGTGT 263  
Db 190103 TTTCTCTTTCTAGAAACAAATATGATTTTATATAAAGTTGATGTTTAAAAAAT 190162  
QY 264 TTATAAGAGAGTAATAAGA 285  
Db 190163 AAAAAATGAAACAAACAAATA 190184

## RESULT 9

US-08-232-463-14/c  
Sequence 14, Application US/08232463  
Patent No. 5670367

GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHEIFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOMLOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-9300  
TELEFAX: (703) 683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

IMMEDIATE SOURCE:  
CLONE: PTZpt-Fls  
US-08-232-463-14

Query Match 7.4%; Score 37.8; DB 2; Length 7218;  
Best Local Similarity 13.5%; Pred. No. 1.1;  
Matches 30; Conservative 103; Mismatches 90; Indels 0; Gaps 0;

QY 148 TATGCAACAAAGAAATGCAATCATCTCTAGTCTCATTTATTTTCATTAGATAG 207  
Db 1494 TAGGATCATCTGTAATTCATATCATATGACAGTGTAAAGATTAAGAAATTTGGTAC 1435  
QY 208 CCGGTTTCTACTACAACTCAATTAAGTGAACGAATGAATGGTTAGTGTAT 267  
Db 1434 RRR 1375  
QY 268 AAAAGAGATTAATAAGTATCATCATCTTTGAGGCAATTAAGGAGAGAGATTACG 327  
Db 1374 RRR 1315  
QY 328 AAACAGTGTCTTACAAAGTGAAACAAAGTTAACTAAAGTGA 370  
Db 1314 RRR 1272

## RESULT 10

US-09-949-016-125489/c  
Sequence 125489, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: C1001307  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 125489  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-125489

Query Match 7.3%; Score 37.6; DB 3; Length 601;  
Best Local Similarity 49.0%; Pred. No. 0.54;  
Matches 100; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 148 TATGCAACAAAGAAATGCAATCATCTCTAGTCTCATTTATTTTCATTAGATAG 207  
Db 365 TATTAATAATATAGTATATCAACAGATGATATATACTGATATAGTAACTACATAT 306  
QY 208 CCGGTTTCTACTACAACTCAATTAAGTGAACGAATGAATGGTTAGTGTAT 267  
Db 305 ACATRTATATAGTGTAGTAACTGTAAGAAAGTGAACAAATTTCTGCTACTATCTTA 246  
QY 268 AAAAGAGATTAATAAGTATCATCATCTTTGAGGCAATTAAGGAGAGAGATTACG 327  
Db 245 TTAATGCAATCTTATTAAGTCTGCAAAATTTTAAGATTTAAATATTTCAAGCTCAGT 186  
QY 328 AAACAGTGTCTTACAAAGTGAA 351  
Db 185 CAACATTAAGCTTAAAGTAAAA 162

## RESULT 11

US-09-949-016-15302/c

; Sequence 15302, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-30  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 15302  
; LENGTH: 44353  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(44353)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-15302

Query Match 7.3%; Score 37.6; DB 3; Length 44353;  
Best Local Similarity 49.0%; Pred. No. 2.5;  
Matches 100; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

Qy 18 TATGCAACAAAGAAATGTCATACATCTCTAGCTCATTTATTTCTATTAGATG 207  
Db 25736 TATTAATAATATGATATATATCAAGCATATATATACATGATATATGTTAACTACATAT 25677  
Qy 208 CCGGTTTTTACTACAACTCAATATAGATGAAAGATGATGCTGTTAT 267  
Db 25676 ACATGATATATAGTTAGTATAGTAAAGATGAAAGATGAAATTTCTGCTACTATCTAA 25617  
Qy 268 AAGAGAGATATTAAGATATCTATCTCTTTGAGCATTAAGGAGAGATTCACG 327  
Db 25616 TTAATGCAATCTTTAATAGTCTCTCAAAATATTAAGAAATTAATTCAAAGCGTCAGT 25557  
Qy 328 AAACAGTGTCTTACAGTGA 351  
Db 25556 CAACATATGCTTAAAGTAAAA 25553

RESULT 12  
US-09-531-120-211  
; Sequence 211, Application US/09531120  
; Patent No. 6972197  
; GENERAL INFORMATION:  
; APPLICANT: PREUSS, DAPHNE  
; APPLICANT: COPENHAVER, GREGORY  
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS  
; FILE REFERENCE: ARD:309US  
; CURRENT APPLICATION NUMBER: US/09/531,120  
; CURRENT FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/125,219  
; PRIOR FILING DATE: 1999-03-18  
; NUMBER OF SEQ ID NOS: 212  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 211  
; LENGTH: 1082144  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-531-120-211

Query Match 7.3%; Score 37.6; DB 4; Length 1082144;  
Best Local Similarity 50.0%; Pred. No. 7.8;  
Matches 94; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

Qy 18 CATGCTTATATGATGAGATAGATATTCTGACCAATTTAGATGCAAAAGCAGAT 77  
Db 431379 CACACATATTAATAAATCATTTTAAAAAATATTTTCAACACAGATTAAGAAATCATTTAA 431438  
Qy 78 TTGTGTTTATCTGTGGTAAATACGTTTCTCCAGTGTATTAAGACCTCCACCG 137  
Db 431439 TTTTATTTTATTTTAAAGTAAACATCTTTATTTCTATTCATTAATGATTCACCA 431498  
Qy 138 TATTAAGTCTTATGCAACAAAGAAATGTCATACATCTCTTAGTCTCATTTATTTT 197  
Db 431499 TAAAAAATATCTGCGGAATATTAATGCGCAAAATTAATAAAAAATTAATTCCTTA 431558  
Qy 198 CATTAGAT 205  
Db 431559 ATTATAT 431566

RESULT 13  
US-08-998-416-187/c  
; Sequence 187, Application US/08998416  
; Patent No. 6239264  
; GENERAL INFORMATION:  
; APPLICANT: Philippen, Peter  
; APPLICANT: Pohlmann, Rainer  
; APPLICANT: Steiner, Sabine  
; APPLICANT: Mohr, Christine  
; APPLICANT: Wendland, Jurgen  
; APPLICANT: Knechtle, Philipp  
; APPLICANT: Redlschning, Corinne  
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPPII  
; NUMBER OF SEQUENCES: 1152  
; CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6239264artis Corporation  
STREET: 3054 Cornwallis Road  
CITY: Research Triangle Park  
STATE: No. 6239264th Carolina  
COUNTRY: USA  
ZIP: 27709  
; COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/998,416  
FILING DATE: 24-DEC-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: CH 0016/97  
FILING DATE: 31-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8587  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 187:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 663 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: PAG1074UP  
US-08-998-416-187

Query Match 7.3%; Score 37.4; DB 3; Length 663;  
Best Local Similarity 58.6%; Pred. No. 0.63;  
Matches 65; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

[illegible]

RESULT 14  
US-08-998-416-779  
; Sequence 779, Application US/08998416

```

1  GENERATED INFORMATION:
2  APPLICANT:  PHILIPSEN, Peter
3  APPLICANT:  Pohlmann, Rainer
4  APPLICANT:  Steiner, Sabine
5  APPLICANT:  Mohr, Christine
6  APPLICANT:  Wendland, Jurgen
7  APPLICANT:  Knechtle, Philipp
8  APPLICANT:  Reischung, Corinne
9  TITLE OF INVENTION:  GENOMIC DNA SEQUENCES OF ASHBYA GOSSEYII
10 TITLE OF INVENTION:  AND USES THEREOF
11 NUMBER OF SEQUENCES:  1152
12
13 CORRESPONDENCE ADDRESS:
14 ADDRESSEE:  No. 6239264artis Corporation
15 STREET:  3054 Cornwallis Road
16 CITY:  Research Triangle Park
17 STATE:  No. 6239264th Carolina
18 COUNTRY:  USA
19 ZIP:  27709
20
21 COMPUTER READABLE FORM:
22 MEDIUM TYPE:  Floppy disk
23
24 COMPUTER:  IBM PC compatible
25 OPERATING SYSTEM:  PC-DOS/MS-DOS
26 SOFTWARE:  Patentin Release #1.0, Version #1.30
27
28 CURRENT APPLICATION DATA:
29 APPLICATION NUMBER:  US/08/998,416
30 FILING DATE:  24-DEC-1997
31
32 CLASSIFICATION:  415
33
34 PRIOR APPLICATION DATA:
35 APPLICATION NUMBER:  CH 0016/97
36 FILING DATE:  31-DEC-1996
37
38 ATTORNEY/AGENT INFORMATION:
39 NAME:  Meigs, J. Timothy
40 REGISTRATION NUMBER:  38,241
41 REFERENCE/DOCKET NUMBER:  Pf/5-30306/A/CGC1976
42
43 TELECOMMUNICATION INFORMATION:
44 TELEPHONE:  919-541-8587
45 TELEFAX:  919-541-8689
46
47 INFORMATION FOR SEQ ID NO:  779:
48
49 SEQUENCE CHARACTERISTICS:
50 LENGTH:  696 base pairs
51 TYPE:  nucleic acid
52 STRANDEDNESS:  single
53 TOPOLOGY:  linear
54
55 MOLECULE TYPE:  DNA (genomic)
56
57 ORIGINAL SOURCE:
58 ORGANISM:  PAG1501RP
59
60 US-08-998-416-779

```

Query Match	7.3%	Score 37.4	DB 3	Length 696
Similarity	58.6%	Pred. No. 0	64	
Best Local	0	Mismatches	46	Indels 0
Matches 65		Conservative		Gaps 0

[illegible]

RESULT 15  
US-08-998-416-1138/c  
; Sequence 1138, Application US/08998416  
; Patent No. 6239264

APPLICANT: Philippse, Peter  
 APPLICANT: Pohlmann, Rainer  
 APPLICANT: Steiner, Sabine  
 APPLICANT: Mohr, Christine  
 APPLICANT: Wendland, Jürgen  
 APPLICANT: Knechtle, Philipp  
 APPLICANT: Redischung, Corinne  
 TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPII  
 TITLE OF INVENTION: AND USES THEREOF  
 NUMBER OF SEQUENCES: 1152  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: No. 623926artis Corporation  
 STREET: 3054 Cornwallis Road  
 CITY: Research Triangle Park  
 STATE: No. 6239264th Carolina  
 COUNTRY: USA  
 REF: 37708

```

COMPUTER READABLE FORM:
MEDIM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.3
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998.416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEO ID NO: 1138:
SEQUENCE CHARACTERISTICS:
LENGTH: 719 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1692UP

```

```
Query Match      Score 37.4 ; DB 3 ; length 719 ;
Similarity      58.6% ;
Beef Local       Pred. No. 0.65 ;
Matches 65 ; Conservative 0 ; Mismatches 46 ; Indels 0 ; Gaps 0 ;
```

QY 180 TATGTCATTATATATTTCTATGATAGCCGGTTTTCCTACAACTCAATTAAGATGAA 239  
Db 574 TACATTTATTTATTAACCTTATTAATGATATAATATATATAA 515  
QY 240 CAGATGAATGGTGTAGTACTCTTTTATAAGAAGATATAAGATACCA 290  
Db 514 GGATATACATTAAATGATATATATAGTTTATTAAGACCAATGAAATACCA 464

Search completed: May 27, 2006, 19:25:41  
Job time : 107.294 secs



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OW nucleic - nucleic search, using sw model

Run on: May 27, 2006, 18:56:27 ; Search time 699.421 Seconds  
(without alignments)  
8994.972 Million cell updates/sec

Title: US-10-099-663-1\_COPY\_1115\_1626

Perfect score: 512  
Sequence: 1 cagtgatccctcattcatcatc.....gtctgcctcagacagaag 512

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA Main:

1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
7: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
8: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
9: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*  
10: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*  
11: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*  
12: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10G\_PUBCOMB.seq:\*  
13: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11A\_PUBCOMB.seq:\*  
14: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11B\_PUBCOMB.seq:\*  
15: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11C\_PUBCOMB.seq:\*  
16: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11D\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	512	100.0	2381	US-10-099-663-1	Sequence 1, Appl
2	326	63.7	336	US-10-099-663-2	Sequence 2, Appl
3	43.6	8.5	413	US-10-674-124A-2159	Sequence 2159, Ap
4	42.8	8.4	558	US-09-925-065A-572321	Sequence 572321,
5	42.8	8.4	558	US-09-925-065A-572321	Sequence 572321,
6	42.8	8.4	679	US-09-925-065A-875384	Sequence 875384,
7	42.8	8.4	679	US-09-925-065A-875384	Sequence 875384,
8	42.8	8.4	679	US-09-925-065A-875384	Sequence 875384,
9	42.8	8.4	679	US-09-925-065A-875384	Sequence 875384,
10	42.8	8.4	679	US-09-925-065A-875384	Sequence 875384,
11	42.8	8.4	679	US-09-925-065A-875384	Sequence 875384,
12	42.8	8.4	900	US-10-301-480-582933	Sequence 582933,
13	42.8	8.4	900	US-10-301-480-582933	Sequence 582933,
14	42.6	8.3	1457619	US-11-098-686-8739	Sequence 8739, Ap
15	42.4	8.3	5908	US-10-239-676-94	Sequence 94, Appl
16	42.4	8.3	5908	US-10-240-453-106	Sequence 106, App
17	42.4	8.3	5908	US-10-221-613-176	Sequence 176, App

C 18	42.2	8.2	1238	4	US-09-925-065A-710652	Sequence 710652,
C 19	42.2	8.2	1238	5	US-09-925-065A-710652	Sequence 710652,
C 20	41.8	8.2	10048	8	US-10-221-613-211	Sequence 211, App
C 21	40.8	8.0	536	12	US-10-301-480-220178	Sequence 220178,
C 22	40.8	8.0	536	12	US-10-301-480-833587	Sequence 833587,
C 23	40.8	8.0	536	12	US-09-925-065A-121976	Sequence 121976,
C 24	40.8	8.0	536	12	US-09-925-065A-121976	Sequence 121976,
C 25	40.6	7.9	7040	6	US-10-172-086-13	Sequence 13, Appl
C 26	40.6	7.9	7040	6	US-10-221-714A-161	Sequence 161, Appl
C 27	40.6	7.9	7040	8	US-10-311-507-47	Sequence 47, Appl
C 28	40.6	7.9	7040	9	US-10-480-846-13	Sequence 13, Appl
C 29	40.6	7.9	7040	9	US-10-473-126-1173	Sequence 173, App
C 30	40.6	7.9	7040	9	US-10-473-126-1173	Sequence 173, App
C 31	40.6	7.9	14095	7	US-10-311-455-450	Sequence 450, App
C 32	40.4	7.9	470	4	US-09-925-065A-178215	Sequence 178215,
C 33	40.4	7.9	470	4	US-09-925-065A-178215	Sequence 178215,
C 34	40.4	7.9	470	5	US-09-925-065A-178215	Sequence 178215,
C 35	40.4	7.9	470	5	US-09-925-065A-178215	Sequence 178215,
C 36	40.4	7.9	471	12	US-10-301-480-269235	Sequence 269235,
C 37	40.4	7.9	471	12	US-10-301-480-269235	Sequence 269235,
C 38	40.4	7.9	471	12	US-10-301-480-882644	Sequence 882644,
C 39	40.4	7.9	471	12	US-10-301-480-882644	Sequence 882644,
C 40	40.4	7.9	1816	4	US-09-925-065A-45365	Sequence 45365, A
C 41	40.4	7.9	1816	5	US-09-925-065A-45365	Sequence 45365, A
C 42	40.4	7.9	1816	12	US-10-301-480-146603	Sequence 146603,
C 43	40.4	7.9	1816	12	US-10-301-480-146603	Sequence 146603,
C 44	40.2	7.9	1400	16	US-11-136-527-6182	Sequence 6182, Ap
C 45	40.2	7.9	1900	16	US-11-136-527-2086	Sequence 2086, Ap

## ALIGNMENTS

RESULT 1  
US-10-099-663-1  
; Sequence 1, Application US/10099663  
; Publication No. US20030177516A1  
; GENERAL INFORMATION:  
; APPLICANT: Avigenics, Inc  
; TITLE OF INVENTION: Avian Gut-Specific Promoters  
; FILE REFERENCE: A181  
; CURRENT APPLICATION NUMBER: US/10/099,663  
; CURRENT FILING DATE: 2002-03-14  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 2381  
; TYPE: DNA  
; ORGANISM: Gallus gallus  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: (1)..(1626)  
; NAME/KEY: exon  
; LOCATION: (1627)..(1693)  
; NAME/KEY: Intron  
; LOCATION: (1694)..(2382)  
; NAME/KEY: exon  
; LOCATION: (2383)..(2381)  
; US-10-099-663-1  
Query Match 100.0%; Score 512; DB 7; Length 2381;  
Best Local Similarity 100.0%; Pred. No. 1.2e-123;  
Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CAGTGATCCCTCATCTCATGCTTATATGAGAGATGATGATTCACCAATTAGA 60  
DB 1115 CAGTGATCCCTCATCTCATGCTTATATGAGAGATGATGATTCACCAATTAGA 1174  
QY 61 ATGGCAAAAGCAGATTTGTTTATCTGTTGGTAAATCGTTTCTCAGTTGAT 120  
DB 1175 ATGGCAAAAGCAGATTTGTTTATCTGTTGGTAAATCGTTTCTCAGTTGAT 1234  
QY 121 AAGAACCTCCACCAAGTAAAGTCTATGCAACAAAGAAATGTCATATCTCTT 180

Db 1235 AAAGACCTCCACAGATTAAGTCTATGCAAGAAAGATTCATTCTCTT 1294  
QY 181 AGTCATATATATTTTCTATTGATAGCCGGTTTTTACTCAACTCAATAGATGAC 240  
Db 1295 AGTCATATATATTTTCTATTGATAGCCGGTTTTTCTCAACTCAATAGATGAC 1354  
QY 241 AGAATGAATGGTTAGTACTGTTTATTAAGAAGTAATTAAGTACTATCATCTTGG 300  
Db 1355 AGAATGAATGGTTAGTACTGTTTATTAAGAAGTAATTAAGTACTATCATCTTGG 1414  
QY 301 AGGCATTAAGGAGGAGAGATTCAGCAACAGTGTGCTTCAAGTGGAAAACAAGTTAA 360  
Db 1415 AGGCATTAAGGAGGAGAGATTCAGCAACAGTGTGCTTCAAGTGGAAAACAAGTTAA 1474  
QY 361 ACTAAGTACCCCTCTCTGACAAATTCACAGTGTGCTTCAAGTGGAAAACAAGTTAA 420  
Db 1475 ACTAAGTACCCCTCTCTGACAAATTCACAGTGTGCTTCAAGTGGAAAACAAGTTAA 1534  
QY 421 TCATCATGTAATTTGCTTCTGATTAAGCCCTGTTCAATTAATCTCTTGGAAAAGCTCTGC 480  
Db 1535 TCATCATGTAATTTGCTTCTGATTAAGCCCTGTTCAATTAATCTCTTGGAAAAGCTCTGC 1594  
QY 481 TACTTACAGAAAGTCTGCTTACAGCAAGAAAG 512  
Db 1595 TACTTACAGAAAGTCTGCTTACAGCAAGAAAG 1626

RESULT 2  
US-10-099-663-2  
; Sequence 2, Application US/10099663  
; Publication No. US20030177516A1  
; GENERAL INFORMATION:  
; APPLICANT: Avigen, Inc  
; TITLE OF INVENTION: Avian Gut-Specific Promoters  
; FILE REFERENCE: A181  
; CURRENT APPLICATION NUMBER: US/10/099,663  
; CURRENT FILING DATE: 2002-03-14  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 336  
; TYPE: DNA  
; ORGANISM: Gallus gallus  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(336)  
; OTHER INFORMATION: chick intestinal fatty acid binding protein promoter region  
US-10-099-663-2

Query Match 63.7%; Score 326; DB 7; Length 336;  
Best Local Similarity 100.0%; Pred. No. 3.9e-75;  
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 181 ATTATTTATTTTCTATGATAGCCGGTTTTTACTCAACTCAATTAAGATGACGAATG 246  
Db 1 ATTATTTATTTTCTATGATAGCCGGTTTTTACTCAACTCAATTAAGATGACGAATG 60  
QY 247 AATGGTTAGTACGCTTTATTAAGAAGTAATTAAGTACTATCATCTTGGAAAAGCAA 306  
Db 61 AATGGTTAGTACGCTTTATTAAGAAGTAATTAAGTACTATCATCTTGGAAAAGCAA 120  
QY 307 TAAAGGAGGAGAGATTCAGCAACAGTGTGCTTCAAGTGGAAAACAAGTTAACTTAA 366  
Db 121 TAAAGGAGGAGAGATTCAGCAACAGTGTGCTTCAAGTGGAAAACAAGTTAACTTAA 180  
QY 367 GTGACCCCTCTCTGACAAATTCACAGTGTGCTTCAAGTGGAAAACAAGTTAACTTAA 426  
Db 181 GTGACCCCTCTCTGACAAATTCACAGTGTGCTTCAAGTGGAAAACAAGTTAACTTAA 240  
QY 427 TGTAAATGCTTCTCTGATTAAGCCCTGTTCAATTAATCTCTTGGAAAAGCTCTGCTTAA 486  
Db 241 TGTAAATGCTTCTCTGATTAAGCCCTGTTCAATTAATCTCTTGGAAAAGCTCTGCTTAA 300

QY 487 CCAGAAAGTCTGCTTACAGCAAGAAAG 512  
Db 301 CCAGAAAGTCTGCTTACAGCAAGAAAG 326

RESULT 3  
US-10-674-124A-2159/c  
; Sequence 2159, Application US/10674124A  
; Publication No. US2004019797A1  
; GENERAL INFORMATION:  
; APPLICANT: INOKO, Hideoshi  
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE  
; FILE REFERENCE: ORIN-003CIP  
; CURRENT APPLICATION NUMBER: US/10/674,124A  
; PRIOR FILING DATE: 2003-09-26  
; PRIOR APPLICATION NUMBER: 10/257,511  
; PRIOR FILING DATE: 2003-03-07  
; PRIOR APPLICATION NUMBER: PCT/JP00/07621  
; PRIOR FILING DATE: 2000-10-30  
; PRIOR APPLICATION NUMBER: JP2000-112699  
; PRIOR FILING DATE: 2000-04-13  
; PRIOR APPLICATION NUMBER: JP2002-327516  
; PRIOR FILING DATE: 2002-09-28  
; PRIOR APPLICATION NUMBER: JP2002-383869  
; PRIOR FILING DATE: 2002-12-09  
; NUMBER OF SEQ ID NOS: 27110  
; SEQ ID NO 2159  
; LENGTH: 413  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Located on chromosome 1  
; FEATURE:  
; OTHER INFORMATION: Distance between a terminus base of telomere on  
; OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base  
; FEATURE:  
; OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of  
; OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and  
; OTHER INFORMATION: 5'-terminus of this base sequence : 205277  
US-10-674-124A-2159

Query Match 8.5%; Score 43.6; DB 9; Length 413;  
Best Local Similarity 55.2%; Pred. No. 0.7;  
Matches 85; Conservative 0; Mismatches 69; Indels 0; Gaps 0;  
QY 138 TATAAGTCTATGCAACAAAGAAATGCAATACATCTCTTATGCTCATTTATTTT 197  
Db 256 TATAGCAACTAGCCAACTAGCTCTCTATTTATTTATTTATTTATTTATTTATTTAT 197  
QY 198 CATTAATAGCCGGTTTTTACTCAACTCAATTAAGATGAAGAAATGAATGAGTTAGT 257  
Db 196 TATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 257  
QY 258 GACTGTTATTAAGAAGCTAATTAAGTACTAT 291  
Db 136 GACTTATTAAGTGAATATGTAATATCTAATTT 103  
RESULT 4  
US-09-925-065A-572321/c  
; Sequence 572321, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135

```

; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 572321
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-572321

```

```

Query Match      8.4%; Score 42.8; DB 4; Length 558;
Best Local Similarity 53.6%; Pred. No. 1.3;
Matches 89; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

```

```

QY 187 ATTATTTTTCATTAGATAGCCGGTTTCTTCACTCAATTAAGATGAAGCAATG 246
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 346 ATTAACTTTTAACTAATTAAGTACCTTTATATCTCTTTAACAAGACATTTAGAACT 287
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 247 AATGGTTAGTACTGTTTATTAAGAAGATTAATTAAGTACTATCATTTGAGCAA 306
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 286 TAAAGTAAAGTCCATGTTAATTAAGAAATTAAGAACTAATTAATTAAGAA 227
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 307 TAAAGGAGGAGAGATTCAGCAACAGTGTCTTCAAGTGAATA 352
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 226 TAAAGCACTTAATCTTGAATTAATTAATTAATTAATTAATTAATTAATTA 181
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 5
US-09-925-065A-572321/C
; Sequence 572321, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 572321
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-572321

```

```

Query Match      8.4%; Score 42.8; DB 5; Length 558;
Best Local Similarity 53.6%; Pred. No. 1.3;
Matches 89; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

```

```

QY 187 ATTATTTTTCATTAGATAGCCGGTTTCTTCACTCAATTAAGATGAAGCAATG 246
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 346 ATTAACTTTTAACTAATTAAGTACCTTTATATCTCTTTAACAAGACATTTAGAACT 287
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 247 AATGGTTAGTACTGTTTATTAAGAAGATTAATTAAGTACTATCATTTGAGCAA 306
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 286 TAAAGTAAAGTCCATGTTAATTAAGAAATTAAGAACTAATTAATTAAGAA 227
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 307 TAAAGGAGGAGAGATTTCAGCAACAGTGTCTTCAAGTGAATA 352
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 226 TAAAGCACTTAATCTTGAATTAATTAATTAATTAATTAATTAATTAATTA 181
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 6
US-09-925-065A-875384/C
; Sequence 875384, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 875384
; LENGTH: 679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-875384

```

```

Query Match      8.4%; Score 42.8; DB 4; Length 679;
Best Local Similarity 53.6%; Pred. No. 1.4;
Matches 89; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

```

```

QY 187 ATTATTTTTCATTAGATAGCCGGTTTCTTCACTCAATTAAGATGAAGCAATG 246
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 515 ATTAACTTTTAACTAATTAAGTACCTTTATATCTCTTTAACAAGACATTTAGAACT 456
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 247 AATGGTTAGTACTGTTTATTAAGAAGATTAATTAAGTACTATCATTTGAGCAA 306
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 455 TAAAGTAAAGTCCATGTTAATTAAGAAATTAAGAACTAATTAATTAAGAA 396
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 307 TAAAGGAGGAGAGATTTCAGCAACAGTGTCTTCAAGTGAATA 352
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 395 TAAAGCACTTAATCTTGAATTAATTAATTAATTAATTAATTAATTAATTA 350
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 7
US-09-925-065A-875385/C
; Sequence 875385, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30

```

```

; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 875385
; LENGTH: 679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-875385
```

```

Query Match      8.4%; Score 42.8; DB 4; Length 679;
Best Local Similarity 53.6%; Pred. No. 1.4;
Matches 89; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
```

```

QY 187 ATTATATTTTCATTAGTAGCCGGTTTTTTTACTACAACTCAATTAAGTGAACAGATG 246
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 515 ATTACATTAAACAATTAGTAGCCCTTTATATACCTTTTACGAAACATTAGAGAACT 456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 247 AATGGGTAGTACTGTTTATTAAGAAGATTAAGATTAAGATTAAGATTAAGATTAAG 306
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 455 TAAAGGTAGTCCATGTATATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 307 TAAAGGAGGAGAGATTTCAGCAACAGTGTCTTACAACTGAGAAA 352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 395 TAAAGACACTAAATCTTAGAAATTAATTAATTAATTAATTAATTAATTAATTA 350
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

RESULT 8
US-09-925-065A-875386/c
; Sequence 875386, Application US/09925065A
; Publication No. US20040181046A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 875386
; LENGTH: 679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-875386
```

```

Query Match      8.4%; Score 42.8; DB 4; Length 679;
Best Local Similarity 53.6%; Pred. No. 1.4;
Matches 89; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
```

```

QY 187 ATTATATTTTCATTAGTAGCCGGTTTTTTTACTACAACTCAATTAAGTGAACAGATG 246
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 515 ATTACATTAAACAATTAGTAGCCCTTTATATACCTTTTACGAAACATTAGAGAACT 456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 247 AATGGGTAGTACTGTTTATTAAGAAGATTAAGATTAAGATTAAGATTAAGATTAAG 306
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 455 TAAAGGTAGTCCATGTATATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 307 TAAAGGAGGAGAGATTTCAGCAACAGTGTCTTACAACTGAGAAA 352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 395 TAAAGACACTAAATCTTAGAAATTAATTAATTAATTAATTAATTAATTAATTA 350
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

RESULT 9
US-09-925-065A-875384/c
; Sequence 875384, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 875384
; LENGTH: 679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-875384
```

```

Query Match      8.4%; Score 42.8; DB 5; Length 679;
Best Local Similarity 53.6%; Pred. No. 1.4;
Matches 89; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
```

```

QY 187 ATTATATTTTCATTAGTAGCCGGTTTTTTTACTACAACTCAATTAAGTGAACAGATG 246
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 515 ATTACATTAAACAATTAGTAGCCCTTTATATACCTTTTACGAAACATTAGAGAACT 456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 247 AATGGGTAGTACTGTTTATTAAGAAGATTAAGATTAAGATTAAGATTAAGATTAAG 306
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 455 TAAAGGTAGTCCATGTATATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 307 TAAAGGAGGAGAGATTTCAGCAACAGTGTCTTACAACTGAGAAA 352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 395 TAAAGACACTAAATCTTAGAAATTAATTAATTAATTAATTAATTAATTAATTA 350
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

RESULT 10
US-09-925-065A-875385/c
; Sequence 875385, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 875385
; LENGTH: 679
```

Page 5

## Polymorph

AAATG	246
AAACT	426

GC/MS 366

```

      APPS      0;
      |
      LATG 246
      |
      LACT 426

```

366 CAA

RESULT 14  
US-11-098-686-8739  
; Sequence 8739, Application US/11098686  
; Publication No. US20060024696A1  
; GENERAL INFORMATION:  
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.  
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES  
; FILE REFERENCE: 09531-128001  
; CURRENT APPLICATION NUMBER: US/11/098,686  
; PRIOR FILING DATE: 2005-04-04  
; PRIOR APPLICATION NUMBER: PCT/US03/31318  
; PRIOR FILING DATE: 2003-10-01  
; PRIOR APPLICATION NUMBER: US 60/416,395  
; PRIOR FILING DATE: 2002-10-04  
; NUMBER OF SEQ ID NOS: 11433  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 8739  
; LENGTH: 1457619  
; TYPE: DNA  
; ORGANISM: Lawsonia intracellularis  
US-11-098-686-8739

Query Match 8.3%; Score 42.6; DB 16; Length 1457619;  
Best Local Similarity 49.6%; Pred. No. 48; Mismatches 164; Indels 6; Gaps 2;  
Matches 167; Conservative 0;

QY 2 AGTGGATCCTTCATCTCATGCTTATTATGAGATGATGATTCGACCAATTAAGA 61  
DB 858309 AGTGAATCTTATCTTATGATCATTTAAATAATATTCATTACAGATGATTA 858368

QY 62 TGGACAAAGCAGAGATTGTGTT--TATCTGTGGGTAATACGTTTCCAGTGT 118  
DB 858369 TATATCAAAAAATTTCTTTTATATTAATGTCACATTAATAATGATTA 858428

QY 119 ATAAAGACCTCCCAAGTATTAAGTCTATGACAAAGAAATGTCATATCATTC 178  
DB 858429 TATATTAACAACATACAAAAAAATTAACAAAAATTTGATTTGACTGTAATTA 858488

QY 179 TTAGTCGATTATTTATTTTCATTAGAT--AGCCGTTTTTTTACTACAACTCAATTAAGA 235  
DB 858489 TTAGCGCTATATATGATATATATATATCTATTAATTAATAGATTAATAGTA 858548

QY 236 TGAACAGATGAATGGTTAGTACTGTTTATTAAGAGATTAATAGATCTATCATC 295  
DB 858549 TAACTATTAATTAAGAAATATTAGTCTATTAATTAATAGAACTAGTTTATGAT 858608

QY 296 ATTGAGCAATTAAGGAGGAGAGATTCAGCAACA 332  
DB 858609 ATAAAGAGGATTAATTTAGAAAGACTATTAACAACA 858645

RESULT 15  
US-10-239-676-94/c  
; Sequence 94, Application US/10239676  
; Publication No. US20030082609A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation  
; FILE REFERENCE: 5013.1003  
; CURRENT APPLICATION NUMBER: US/10/239,676  
; CURRENT FILING DATE: 2002-09-24  
; PRIOR APPLICATION NUMBER: PCT/EP01/03968  
; DE 10019058.8  
; DE 10019173.8  
; DE 10032529.7  
; DE 10043826.1  
; PRIOR FILING DATE: 2001-04-06  
; 2000-04-06  
; 2000-04-07

2000-06-30  
; 2000-09-01  
; NUMBER OF SEQ ID NOS: 228  
; SEQ ID NO 94  
; LENGTH: 5908  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-239-676-94

Query Match 8.3%; Score 42.4; DB 6; Length 5908;  
Best Local Similarity 54.5%; Pred. No. 4.7;  
Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 138 TATAAGTCCTATGCAACAAGAAATGCAATCATTCCTTATGTCATTATTTAT 197  
DB 4703 TATACTCTTTTATTAATAAACAATATATATATATATATATATATATAT 4644

QY 198 CATTAGATAGCCGGTTTTTTTACTACAACTCAATATAGATGACAGATGAATGGTTAGT 257  
DB 4643 TATTAATTTTATTTTCACTACTCTTCTCTAATAAATTAATAATCTATACCTAT 4584

QY 258 GACTGTTTATTAAGAGATTAATTAAGATCTATCA 293  
DB 4583 TCCCTTTATCCAAATATATATCTATATACGAACA 4548

Search completed: May 27, 2006, 20:03:05  
Job time : 700.421 secs

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OW nucleic - nucleic search, using BW model

Run on: May 27, 2006, 18:57:23 ; Search time 9.209 Seconds  
(without alignments)  
6280.280 Million cell updates/sec

Title: US-10-099-663-1\_COPY\_1115\_1626  
Sequence: 1 cagcggatccctcattccat.....gtctgcctcagacagaaag 512

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 242596 seqs, 56442199 residues

Total number of hits satisfying chosen parameters: 485192

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA\_New:\*  
1: /EMC\_Celestra\_SIDS3/ptodata/1/pubpna/US09\_NEW\_PUB\_seq.\*  
2: /EMC\_Celestra\_SIDS3/ptodata/1/pubpna/US06\_NEW\_PUB\_seq.\*  
3: /EMC\_Celestra\_SIDS3/ptodata/1/pubpna/US07\_NEW\_PUB\_seq.\*  
4: /EMC\_Celestra\_SIDS3/ptodata/1/pubpna/US08\_NEW\_PUB\_seq.\*  
5: /EMC\_Celestra\_SIDS3/ptodata/1/pubpna/PCT\_NEW\_PUB\_seq.\*  
6: /EMC\_Celestra\_SIDS3/ptodata/1/pubpna/US10\_NEW\_PUB\_seq.\*  
7: /EMC\_Celestra\_SIDS3/ptodata/1/pubpna/US11\_NEW\_PUB\_seq.\*  
8: /EMC\_Celestra\_SIDS3/ptodata/1/pubpna/US60\_NEW\_PUB\_seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	33.6	6.6	1638	US-11-217-529-79683	Sequence 79683, A
C 2	33	6.4	1997	US-11-293-697-546	Sequence 546, App
C 3	32.8	6.4	2873	US-11-293-697-727	Sequence 727, App
C 4	32.6	6.4	995	US-11-233-726-22	Sequence 22, App1
C 5	32.6	6.4	1616	US-10-953-349-2269	Sequence 2269, App
C 6	32.6	6.4	5204	US-10-524-021-12	Sequence 12, App1
C 7	32.4	6.3	1221	US-11-217-529-76807	Sequence 76807, A
C 8	32.4	6.3	3064	US-10-953-349-24676	Sequence 24676, A
C 9	32	6.2	670	US-11-301-554-1289	Sequence 1289, App
C 10	31.6	6.2	2276	US-10-953-349-3916	Sequence 3916, App
C 11	31.2	6.1	135090	US-10-505-928-607	Sequence 607, App
C 12	31	6.1	600	US-11-217-529-81653	Sequence 81653, A
C 13	30.6	6.0	1301	US-10-953-349-1126	Sequence 1126, App
C 14	30.6	6.0	1503	US-11-217-529-1697	Sequence 1697, App
C 15	30.4	5.9	453	US-11-217-529-5137	Sequence 5137, App
C 16	30.4	5.9	1524	US-11-217-529-77900	Sequence 77900, App
C 17	30.4	5.9	2235	US-11-217-529-82000	Sequence 82000, App
C 18	30.4	5.9	8372	US-11-301-554-1893	Sequence 1893, App
C 19	30.2	5.9	1893	US-11-217-529-77169	Sequence 77169, App
C 20	30.2	5.9	138941	US-10-489-730-10	GENERAL INFORMATION
C 21	29.8	5.8	228	US-11-217-529-174019	Sequence 174019, App
C 22	29.8	5.8	1506	US-11-217-529-2930	Sequence 2930, App
C 23	29.8	5.8	1725	US-11-217-529-75449	Sequence 75449, App
C 24	29.8	5.8	2343	US-11-217-529-1859	Sequence 1859, App
C 25	29.6	5.8	1383	US-10-953-349-12683	Sequence 12683, A

C 26	29.6	5.8	2149	US-10-953-349-19029	Sequence 19029, A
C 27	29.6	5.8	2830	US-11-293-697-1209	Sequence 1209, App
C 28	29.6	5.8	3408	US-10-505-928-184	Sequence 184, App
C 29	29.6	5.8	3408	US-10-511-937-337	Sequence 337, App
C 30	29.6	5.8	135090	US-10-505-928-607	Sequence 607, App
C 31	29.4	5.7	1329	US-11-217-529-76309	Sequence 76309, A
C 32	29.4	5.7	1391	US-10-953-349-8646	Sequence 8646, App
C 33	29.4	5.7	1592	US-10-953-349-22352	Sequence 22352, A
C 34	29.2	5.7	4050	US-11-217-529-6036	Sequence 6036, App
C 35	29.2	5.7	544	US-10-953-349-14832	Sequence 14832, A
C 36	29	5.7	864	US-10-953-349-26473	Sequence 26473, A
C 37	29	5.7	958	US-10-953-349-6057	Sequence 6057, App
C 38	29	5.7	1059	US-11-217-529-80017	Sequence 80017, A
C 39	29	5.7	2829	US-11-217-529-3101	Sequence 3101, App
C 40	29	5.7	5849	US-11-270-287-1	Sequence 1, App1
C 41	29	5.7	22118	US-11-284-877-16	Sequence 16, App1
C 42	28.8	5.6	512	US-11-301-554-1707	Sequence 1707, App
C 43	28.8	5.6	779	US-11-301-554-1875	Sequence 187, App
C 44	28.8	5.6	841	US-10-953-349-8745	Sequence 8745, App
C 45	28.8	5.6	1173	US-11-217-529-4261	Sequence 4261, App

## ALIGNMENTS

RESULT 1  
US-11-217-529-79683/c  
Sequence 79683, Application US/11217529  
Publication No. US20060099612A1  
GENERAL INFORMATION:  
APPLICANT: SUNTORY LIMITED  
APPLICANT: NAKAO, YOSHIHIRO  
APPLICANT: NAKAMURA, NORIHISA  
APPLICANT: KODAMA, YUKIKO  
APPLICANT: FUJIMURA, TOMOKO  
APPLICANT: ASHIKARI, TOSHIHIKO  
TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS  
FILE REFERENCE: S-38-285  
CURRENT APPLICATION NUMBER: US/11/217,529  
CURRENT FILING DATE: 2005-09-02  
PRIOR APPLICATION NUMBER: US 10/932,182  
PRIOR FILING DATE: 2004-09-02  
NUMBER OF SEQ ID NOS: 197023  
SOFTWARE: Patent in version 3.3  
SEQ ID NO 79683  
LENGTH: 1638  
TYPE: DNA  
ORGANISM: Saccharomyces pastorianus  
US-11-217-529-79683  
Query Match 6.6%; Score 33.6; DB 7; Length 1638;  
Best Local Similarity 52.9%; Pred. No. 1.4;  
Matches 72; Conservative 0; Mismatches 64; Indels 0; Gaps 0;  
QY 221 TAAGTGAACAGAAATGAGTGTGCTTTATTAAGAAGTATAAGATCTA 290  
DB 304 TTAATGACACGACATGAAATGCTTGCAATTTGGAGAGAGAAATCAATTAAT 245  
QY 291 TCATCATTTAGGACATTAAGGAGAGAGATTCAGCAACAGTGTCTTACAGTGA 350  
DB 244 TTTTGTTTTCAACCAATTTGGCAACCATATCTCTTCTGTTAAGATATCTGGAT 185  
QY 351 AACAGTTAACTAA 366  
DB 184 GACAGTTTGTGTA 169  
RESULT 2  
US-11-293-697-546/c  
Sequence 546, Application US/11293697  
Publication No. US20060105376A1  
GENERAL INFORMATION:  
APPLICANT: HELIX RESEARCH INSTITUTE

```
/ TITLE OF INVENTION: Novel full length cDNA
/ FILE REFERENCE: H1-A0106
/ CURRENT APPLICATION NUMBER: US/11/293,697
/ CURRENT FILING DATE: 2005-12-05
/ PRIOR APPLICATION NUMBER: US/10/108,260
/ PRIOR FILING DATE: 2002-03-28
/ NUMBER OF SEQ ID NOS: 5458
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 546
/ LENGTH: 1997
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-11-293-697-546
```

```
Query Match          6.4%; Score 33; DB 7; Length 1997;
Best Local Similarity 52.6%; Pred. No. 2.2;
Matches 72; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
```

```
QY 153 AACAAAGAAATGTCATTCATTCTTCTAGTCATTTATTTTCATTAGATAGCCGGT 212
DB 1759 AATAGAGATTAGTCAAAATATTTCTGTTAAAAATTTAAATTCAGATGAGAAAT 1700
QY 213 TTTTACTACAACTCAATTAAGTGAACGAATGAATGGGTTAGTACTTTTATTAAGA 272
DB 1699 TTATGCATTAATAATCATATTTGATATATCAGATGAATCTTCAGTGTCTTTCTGAAAA 1640
QY 273 AGAGATATAAGATACT 289
DB 1639 CAATATCTACTTTAGT 1623
```

```
RESULT 3
US-11-293-697-727/c
/ Sequence 727, Application US/11293697
/ Publication No. US20060105376A1
/ GENERAL INFORMATION:
/ APPLICANT: HELIX RESEARCH INSTITUTE
/ TITLE OF INVENTION: Novel full length cDNA
/ FILE REFERENCE: H1-A0106
/ CURRENT APPLICATION NUMBER: US/11/293,697
/ CURRENT FILING DATE: 2005-12-05
/ PRIOR APPLICATION NUMBER: US/10/108,260
/ PRIOR FILING DATE: 2002-03-28
/ NUMBER OF SEQ ID NOS: 5458
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 727
/ LENGTH: 2873
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-11-293-697-727
```

```
Query Match          6.4%; Score 32.8; DB 7; Length 2873;
Best Local Similarity 59.8%; Pred. No. 2.9;
Matches 55; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
```

```
QY 154 ACAAGAAGAAATGTCATTCATTCTTCTAGTCATTTATTTTCATTAGATAGCCGGT 213
DB 1468 AAAATTAATTGAGCTTAACATATTTAATTTTGAATAATTAGACATTAATTTGTTCT 1409
QY 214 TTTTACTACAACTCAATTAAGTGAACGAAT 245
DB 1408 TTAATTTACCACTCAGTTATAGTGAAGAAAT 1377
```

```
RESULT 4
US-11-233-726-22
/ Sequence 22, Application US/11233726
/ Publication No. US20060090216A1
/ GENERAL INFORMATION:
/ APPLICANT: APURA, Nestor
/ APPLICANT: KNOX, Shing
/ APPLICANT: ALEXANDROV, Nikolai
/ APPLICANT: TATARINOVA, Tatiana
```

```
/ APPLICANT: FANG, Yiwen
/ APPLICANT: PENNELL, Roger
/ APPLICANT: LU, Yu-Ping
/ APPLICANT: MEDRANO, Leonard
/ APPLICANT: COOK, Zhong A.
/ APPLICANT: FELDMAN, Kenneth C.
/ TITLE OF INVENTION: Promoter, Promoter Control Elements, And Uses Th
/ FILE REFERENCE: 2750-1619PUS2
/ CURRENT APPLICATION NUMBER: US/11/233,726
/ CURRENT FILING DATE: 2005-02-23
/ PRIOR APPLICATION NUMBER: 60/612,891
/ PRIOR FILING DATE: 2004-09-23
/ PRIOR APPLICATION NUMBER: 60/613,134
/ PRIOR FILING DATE: 2004-09-23
/ PRIOR APPLICATION NUMBER: 60/637,174
/ PRIOR FILING DATE: 2004-12-16
/ NUMBER OF SEQ ID NOS: 46
/ SEQ ID NO 22
/ LENGTH: 995
/ TYPE: DNA
/ ORGANISM: Arabidopsis thaliana
```

```
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (1009)..(1009)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (1)..(995)
/ OTHER INFORMATION: Cereas Promoter construct YP0385 as found in Promoter Report #146
US-11-233-726-22
```

```
Query Match          6.4%; Score 32.6; DB 7; Length 995;
Best Local Similarity 55.9%; Pred. No. 2.3;
Matches 62; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
```

```
QY 172 CATTCTTACTGTCATTTATTTTCATTAGATAGCCGGTTTTTACTACAACTCAAT 231
DB 237 CATGCCATATACACAATTTTGTTCGTCGATGATGATGGTTATATATTTCTAAA 296
QY 232 AAGATGAACAGATGAATGGGTTAGTGAATGTTTATTAAGAAGTAAATA 282
DB 297 AAACTAACAATTTTACTGATTTGAAGTGAATATTTGACACTAATATTA 347
```

```
RESULT 5
US-10-953-349-2269/c
/ Sequence 2269, Application US/10953349
/ Publication No. US20060107345A1
/ GENERAL INFORMATION:
/ APPLICANT: ALEXANDROV, Nikolai et al.
/ TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
/ FILE REFERENCE: 2750-1579PUS2
/ CURRENT APPLICATION NUMBER: US/10/953,349
/ CURRENT FILING DATE: 2004-09-30
/ NUMBER OF SEQ ID NOS: 40252
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 2269
/ LENGTH: 1616
/ TYPE: DNA
/ ORGANISM: Arabidopsis thaliana
US-10-953-349-2269
```

```
Query Match          6.4%; Score 32.6; DB 6; Length 1616;
Best Local Similarity 58.9%; Pred. No. 2.7;
Matches 56; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
```

```
QY 107 TTCTCCAGTTGATTAAGACCCCTCCACAGATTAAGTCTTAGCAACAAGAAATGT 166
DB 1598 TTCTTTCTTCTTAAGATGATGACAGTACTATACAGAGAAATCCACAATGAGAAAA 1539
QY 167 CAATACATTCCTTAGTCTCATTTATTTTCAAT 201
```



Db 1538 CAAAAAAGTAAATCTACTTTTCTTTCTTT 1504

## RESULT 6

US-10-524-021-12  
; Sequence 12, Application US/10524021  
; Publication No. US2006009590A1

## GENERAL INFORMATION:

APPLICANT: NAGOYA INDUSTRIAL SCIENCE RESEARCH INSTITUTE  
APPLICANT: GIFU INTERNATIONAL INSTITUTE OF BIOTECHNOLOGY

APPLICANT: YAMADA, Yoshiji

APPLICANT: YOKOTA, Mitsuhito

TITLE OF INVENTION: Method for diagnosing a risk of restenosis after percutaneous cor

FILE REFERENCE: C0200501

CURRENT APPLICATION NUMBER: US/10/524,021

PRIOR FILING DATE: 2005-02-09

PRIOR APPLICATION NUMBER: JP P2002-233041

NUMBER OF SEQ ID NOS: 67

SOFTWARE: PatentIn version 3.1

SEQ ID NO 12

LENGTH: 5204

TYPE: DNA

ORGANISM: Homo sapiens

US-10-524-021-12

Query Match 6.4%; Score 32.6; DB 6; Length 5204;

Best Local Similarity 58.9%; Pred. No. 4;

Matches 56; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 138 TATAAAGTCTTATGCAACAGAAATGTCATATCTCTAGTCTCATTTATTTT 197

Db 2987 TAGAAGCATCATTTTCAACAAATTTTAAACCTTTTGTAGCTCAACATTTTCT 3046

Qy 198 CATTAATGACCGGTTTCTTACTACACTCAATA 232

Db 3047 ATTAAATTACATGTTGTGAATGACATTTTA 3081

Db 3047 ATTAAATTACATGTTGTGAATGACATTTTA 3081

## RESULT 7

US-11-217-529-76807/c

; Sequence 76807, Application US/11217529

; Publication No. US2006009612A1

## GENERAL INFORMATION:

APPLICANT: SUNTORY LIMITED

APPLICANT: NAKAO, YOSHIHIRO

APPLICANT: NAKAMURA, NORIHISA

APPLICANT: KODAMA, YUKIKO

APPLICANT: FUJIMURA, TOMOKO

APPLICANT: ASHIKARI, TOSHIHIKO

TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS

FILE REFERENCE: S-38-285

CURRENT APPLICATION NUMBER: US/11/217,529

PRIOR FILING DATE: 2005-09-02

PRIOR APPLICATION NUMBER: US 10/932,182

NUMBER OF SEQ ID NOS: 197023

SOFTWARE: PatentIn version 3.3

SEQ ID NO 76807

LENGTH: 1221

TYPE: DNA

ORGANISM: Saccharomyces pastorianus

US-11-217-529-76807

Query Match 6.3%; Score 32.4; DB 7; Length 1221;

Best Local Similarity 57.2%; Pred. No. 2.8;

Matches 79; Conservative 0; Mismatches 56; Indels 3; Gaps 1;

Qy 168 AATATATCTTATGTCATATATTTTATGATGCGGTTTCTTACTACATC 227

Db 601 AATATATCTTATGTCATATATTTTATGATGCGGTCATTTGTTCAATTCCTTTTCT 542

Qy 228 AATATAGTAAACAGATGATGGTTAGTACTGTTAT---AAGAGAGTAAAG 284

Db 541 ACTTCTCAAGACAGATCAAGCGCTTACAGTGTACTCTATATAACAGATAG 482

Qy 285 ATACTATCATCTTTGAG 302

Db 481 ATCTATCATTTATTTAG 464

## RESULT 8

US-10-953-349-24676

; Sequence 24676, Application US/10953349

; Publication No. US20060107345A1

## GENERAL INFORMATION:

APPLICANT: ALEXANDROV, Nikolai et al.

TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES

FILE REFERENCE: 2750-1579PUS2

CURRENT APPLICATION NUMBER: US/10/953,349

PRIOR FILING DATE: 2004-09-30

NUMBER OF SEQ ID NOS: 40252

SOFTWARE: PatentIn version 3.3

SEQ ID NO 24676

LENGTH: 3064

TYPE: DNA

ORGANISM: Glycine max

US-10-953-349-24676

Query Match 6.3%; Score 32.4; DB 6; Length 3064;

Best Local Similarity 50.6%; Pred. No. 3.8;

Matches 78; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

Qy 219 CTACAATCAATTAAGTAAAGACAGATGATGGTTAGTACTGTTATTAAGAGTA 278

Db 2841 CTAGTAGGCGATGAAATGAATTAATGATTAATTTTGGCTTTGTAAGAGAA 2900

Qy 279 ATTAAGATATCTATCATCTTTGAGGCAATTAAGAGGAGAGATTCAGCAACATGTC 338

Db 2901 AAAAAGAAAGAAAGAAAGCAAGCTGGATTTAAATAATTTCTAGCTCACCGATGGGGGTG 2960

Qy 339 TTACAAGTGAAGAAAGCAAGTTAACTTAAGTGACC 372

Db 2961 TTGAACCTTGAAATGATTTGTGATTAAGATCTCC 2994

## RESULT 9

US-11-301-554-1289

; Sequence 1289, Application US/11301554

; Publication No. US20060088527A1

## GENERAL INFORMATION:

APPLICANT: Henderson, Robert A.

APPLICANT: Wang, Tonglong

APPLICANT: Watanabe, Yoshihiro

APPLICANT: Kalos, Michael D.

APPLICANT: Sleath, Paul R.

APPLICANT: Johnson, Jeffrey C.

APPLICANT: Retter, Marc W.

APPLICANT: Durham, Margarita

APPLICANT: Carter, Derrick

APPLICANT: Fanger, Gary R.

APPLICANT: Vedvick, Thomas S.

APPLICANT: Bangur, Chaitanya S.

APPLICANT: McNabb, Andria

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

FILE REFERENCE: 210121.478C21

CURRENT APPLICATION NUMBER: US/11/301,554

PRIOR FILING DATE: 2005-12-13

PRIOR APPLICATION NUMBER: US 10/283,017

PRIOR FILING DATE: 2002-10-28

PRIOR APPLICATION NUMBER: US 10/113,872

PRIOR FILING DATE: 2002-03-28

PRIOR APPLICATION NUMBER: US 10/017,754

```

; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 09/902,941
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 09/849,626
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: US 09/736,457
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 09/702,705
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: US 09/677,419
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 09/671,325
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US 09/658,824
; PRIOR FILING DATE: 2000-09-08
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2157
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 1289
; LENGTH: 670
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-301-554-1289
```

```

Query Match          6.2%; Score 32; DB 7; Length 670;
Best Local Similarity 45.8%; Pred. No. 2.9;
Matches 110; Conservative 0; Mismatches 130; Indels 0; Gaps 0;
```

```

QY 46 TCTCACCATAATGAGATGACCAAGAGATTGTTGTTTCTGTTGGTAAATACGT 105
DB 108 TTTCAATATAAATAAGAGGTAAATGTTAAATACGTATTCATGTTCAATACATTTAT 167
QY 106 TTTTCCAGTTGTATTAAGACCCCTCCACAGTATTAAGTCTATGCAACAAAGAAATG 165
DB 168 CTGAAGAAGTTATTAATAAACAACATGTAAGCTCTGATTTAGGGAAGAAATTCATT 227
QY 166 TCAATACATCTCTAGTCTCATTTATTTTCAATTAGATGACCGGTTTCTACAAAC 225
DB 228 TTTGTAAATTTTCATAGATTAAATTTTACACAGAACTTATTCATAGTTAGATGCA 287
QY 226 TCAATAGATGAACAGATGAATGGGTAGTGAAGCTGTTTAAAGAAAGATTAAGA 285
DB 288 TTAGGTGCAAACTTCCAAAGAAAGGTTAGTGTATTAAAGAAACGTCATTAACA 347

RESULT 10
US-10-953-349-3916/c
; Sequence 3916, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 3916
; LENGTH: 2276
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-3916
```

```

Query Match          6.2%; Score 31.6; DB 6; Length 2276;
Best Local Similarity 58.5%; Pred. No. 5.7;
Matches 55; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
```

```

QY 196 TTCATTAGATGCGCGTTTCTACTACAATCAATAAGATGAACAGATGAATGGTTA 255
DB 2164 TTCCAAAGAAACCGTCTTCATAATTAATAAACAACAAACACACATGATGATGAACA 2105
QY 256 GTGACTGTTTATAAAGAAAGATTAATAAGATACT 289
```

```

DB 2104 TTGATTGTGATGATGATGTTAATGATGATGAT 2071
```

```

RESULT 11
US-10-505-928-607
; Sequence 607, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO: 607
; LENGTH: 135090
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-505-928-607
```

```

Query Match          6.1%; Score 31.2; DB 6; Length 135090;
Best Local Similarity 51.4%; Pred. No. 30;
Matches 72; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
```

```

QY 113 AGTTGTAAAGACCTCCGACAGATTAAGTCTATGCAACAAAGAAATGTCATAC 172
DB 19941 AGATGTTTAAATGTTCAAGCCCTCTGTAAGCAGATATTAAATAATTAATTTTAG 20000
QY 173 ATTCCTAGCTCATATATTTTCAATTAGATAGCCGGTTTCTACAACTCAATA 232
DB 20001 ATTTGCTCTTCATATTTTCTCATTTAAATAAACCCTGTGAACCCCATCTACTA 20060
QY 233 AGATGAACAGATGAATGGG 252
DB 20061 AAATTAACAATAATGCTGGG 20080
```

```

RESULT 12
US-11-217-529-81653
; Sequence 81653, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIRO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 81653
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-81653
```

```

Query Match          6.1%; Score 31; DB 7; Length 600;
Best Local Similarity 59.8%; Pred. No. 5.3;
Matches 52; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```

QY 169 ATACATTCCTTAGTGTCTATTAATTTTCAATTAGATACCGGTTTCTACAACTCA 228
DB 125 ATACCTATTCCTGTCACAAATTTCTTATCTTTTCAAGAAATATTTTCAACAAATACA 184
```

Qy 229 AATAAGATGACGAATGATGGTTA 255  
Db 185 AGTACATTAATATATTAATACGGTA 211

## RESULT 13

US-10-953-349-1126/c  
; Sequence 1126, Application US/10953349  
; Publication No. US20060107345A1  
; GENERAL INFORMATION:  
; APPLICANT: ALEXANDROV, Nikolai et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1579PUS2  
; CURRENT APPLICATION NUMBER: US/10/953,349  
; PRIOR FILING DATE: 2004-09-30  
; NUMBER OF SEQ ID NOS: 40252  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 1126  
; LENGTH: 1301  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-10-953-349-1126

Query Match 6.0%; Score 30.6; DB 6; Length 1301;  
Best Local Similarity 49.1%; Pred. No. 8.9;  
Matches 81; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

Qy 142 AAGCTCTATGCAACAAAGAAATGCAATATCATTCCTAGTCTCATTTATTTTCATT 201  
Db 1233 AATTTGTAATAAGATATGCGCAATATCAAACTCTCTCTCTCTCATGTTTATG 1174  
Qy 202 AGATAGCCGCTTTTACTACCACTCAATTAAGTAAGAATGAATGGTTAGTACT 261  
Db 1173 GTPAATGATCTTTTGATCCGAGCCAAAGCTGAGTCAAGAGTGGCGGTTGAGCC 1114  
Qy 262 GTTTTAAAGAAAGATTAAGATATCATTCATTCATTGAGCAA 306  
Db 1113 ATGATATAGGCGCGCATATGATGTTGTTCCGCCAATAAGCA 1069

## RESULT 14

US-11-217-529-1697  
; Sequence 1697, Application US/11217529  
; Publication No. US20060099612A1  
; GENERAL INFORMATION:  
; APPLICANT: SUNTORY LIMITED  
; APPLICANT: NAKAO, YOSHIHIRO  
; APPLICANT: NAKAMURA, NORIHISA  
; APPLICANT: KODAMA, YUKIKO  
; APPLICANT: FUJIMURA, TOMOKO  
; APPLICANT: ASHIKARI, TOSHIHIRO  
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS  
; FILE REFERENCE: S-38-285  
; CURRENT APPLICATION NUMBER: US/11/217,529  
; PRIOR FILING DATE: 2005-09-02  
; PRIOR APPLICATION NUMBER: US 10/932,182  
; NUMBER OF SEQ ID NOS: 197023  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 1697  
; LENGTH: 1503  
; TYPE: DNA  
; ORGANISM: Saccharomyces pastorianus  
US-11-217-529-1697

Query Match 6.0%; Score 30.6; DB 7; Length 1503;  
Best Local Similarity 52.8%; Pred. No. 9.3;  
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

Qy 44 ATTCTACCAATTAATGATGACAAAGAGATTTGTTTATCTGTTGGTAATAC 103

Db 570 AATTCTCAACATTTCTATTAATAAACAAGATTGGATCTTTCTGTCATCTTC 629  
Qy 104 GTTTCTCCAGTTGTATAAGACCTCCACAGATTAAGTCTTATGCAACAAAGAAA 163  
Db 630 TAAATTAACCTTGATTAAGAACCCCATGATATGCCCAACAACTTTAAATTAACA 669  
Qy 164 TGTCA 168  
Db 690 TGTCA 694

## RESULT 15

US-11-217-529-5137  
; Sequence 5137, Application US/11217529  
; Publication No. US20060099612A1  
; GENERAL INFORMATION:  
; APPLICANT: SUNTORY LIMITED  
; APPLICANT: NAKAO, YOSHIHIRO  
; APPLICANT: NAKAMURA, NORIHISA  
; APPLICANT: KODAMA, YUKIKO  
; APPLICANT: FUJIMURA, TOMOKO  
; APPLICANT: ASHIKARI, TOSHIHIRO  
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS  
; FILE REFERENCE: S-38-285  
; CURRENT APPLICATION NUMBER: US/11/217,529  
; PRIOR FILING DATE: 2005-09-02  
; PRIOR APPLICATION NUMBER: US 10/932,182  
; NUMBER OF SEQ ID NOS: 197023  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 5137  
; LENGTH: 453  
; TYPE: DNA  
; ORGANISM: Saccharomyces pastorianus  
US-11-217-529-5137

Query Match 5.9%; Score 30.4; DB 7; Length 453;  
Best Local Similarity 46.0%; Pred. No. 7;  
Matches 103; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

Qy 147 CTATGCAACAAAGAAATGCAATATCATTCCTTAGTCTCATTTATTTTCATTGATA 206  
Db 8 CTGCCCAAGAAAGTAATCAAGCTTGAACTCAACCCAGAGGTCCTTTAAAGAAA 67  
Qy 207 GCCGTTTTTACTCAACTCAATTAAGATGAACGAATGAGGTGAGCTGTTTA 266  
Db 68 GAAGAATGCTATAGACTTGAATCGAGACAAAGACTTACCAAGAAAGAAAGGAAG 127  
Qy 267 TAAAGAAGTAATAAGATCTATCATTCATTTGAGCAATTAAGGAGGAGATTCAG 326  
Db 128 AACAAATCAAGAAAGAAAGAACTTACCAAAACAGATTGTCTAGAGCAGAGCATAGTGC 187  
Qy 327 CAAACAGTGTCTTACAAAGTGAAGAAACAGTTAACTAAAGTGA 370  
Db 188 CCAAGACTTGGCTACATCCAGAGAAAGAAAGAAATTAAGAGA 231

Search completed: May 27, 2006, 20:04:03  
Job time : 10.2029 secs

**THIS PAGE BLANK (USPTO)**